```
                GenCore version 5.1.6
           Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame plus_p2n model

Run on:       October 24, 2004, 03:55:11 ; Search time 2708 seconds
                                            (without alignments)
                                            1776.235 Million cell updates/sec

Title:         US-10-613-105-2
Perfect score: 762
Sequence:      1 MKSSGLPPFLVLALGTLAP.........RDLKCCMGCGKSCVSPVKA 132

Scoring table: BLOSUM62
               Xgapop 10.0 , Xgapext  0.5
               Ygapop 10.0 , Ygapext  0.5
               Fgapop  6.0 , Fgapext  7.0
               Delop   6.0 , Delext   7.0

Searched:      32882875 seqs, 1821865908 residues

Total number of hits satisfying chosen parameters:       65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 100 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US10613105/runat_22102004_170441_18696/app_query.fasta_1.327
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=45 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10613105.@CGN_1_1_3437 @runat_22102004_170441_18696 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :     EST:*
               1: .gb_est1:*
               2: .gb_est2:*
               3: .gb_htc:*
               4: .gb_est3:*
               5: .gb_est4:*
               6: .gb_est5:*
               7: .gb_est6:*
               8: .gb_gss1:*
               9: .gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	762	100.0	477	1	AA460433	AA460433 zx66h04.r
c 2	762	100.0	534	4	BI490742	BI490742 603032084
3	762	100.0	535	4	BM769596	BM769596 K-EST0052
4	762	100.0	543	6	CD690842	CD690842 EST7365 h
5	762	100.0	554	6	CD707973	CD707973 EST24500
6	762	100.0	563	5	BX089446	BX089446 BX089446
7	762	100.0	583	4	CK430164	CK430164 oj45h07.y
8	762	100.0	593	4	BM767014	BM767014 K-EST0049
9	762	100.0	597	4	BM843274	BM843274 K-EST0121
10	599	100.0	762	6	CD000587	CD000587 AGENCOURT
11	603	100.0	762	4	CK429948	CK429948 oj43c03.y
c 12	606	100.0	762	5	BM842052	BM842052 K-EST0119
13	612	100.0	762	6	BM976429	BM976429 UI-CF-EN1
14	626	100.0	762	6	BQ232871	BQ232871 AGENCOURT
15	760	100.0	762	6	CD103997	CD103997 AGENCOURT
16	796	100.0	762	9	CD000527	CD000527 AGENCOURT
17	874	99.5	758	7	BG533465	BG533465 601860869
18	394	99.5	758	6	AY417768	AY417768 Homo sapi
c 19	597	99.5	758	7	CF529326	CF529326 UI-1-BC1p
c 20	598	99.5	758	6	BU753561	BU753561 UI-1-BC1-
c 21	599	99.5	758	5	CB305723	CB305723 UI-CF-EN1
c 22	607	99.5	758	6	CF529317	CF529317 UI-1-BC1p
23	623	99.3	757	5	CD685198	CD685198 EST1718 h
24	524	99.3	757	6	CD689844	CD689844 EST6367 h
25	589	99.3	757	5	BQ436671	BQ436671 AGENCOURT
26	608	99.3	757	6	BI253834	BI253834 60294607
27	610	99.3	757	6	CD108133	CD108133 AGENCOURT
28	612	99.3	757	6	CB152458	CB152458 K-EST0209
29	906	99.3	757	6	BQ691723	BQ691723 AGENCOURT
30	541	99.2	756	2	CD690425	CD690425 EST6948 h
31	582	99.2	756	2	AA572950	AA572950 nm33a03.8
32	582	99.2	754	2	CD690400	CD690400 EST6923 h
33	633	99.0	754	4	BU927919	BU927919 AGENCOURT
34	493	98.2	748	4	BM694343	BM694343 UI-E-CI1-
35	567	98.2	748	4	BM704851	BM704851 UI-E-CI1-
36	962	98.2	748	6	BI489624	BI489624 603032084
37	575	98.0	747	4	AW264225	AW264225 xq99c06.x
c 38	580	97.9	746	6	CB151317	CB151317 K-EST0208
39	483	97.1	740	2	BE396187	BE396187 601308817
40	581	96.9	738	6	BE392080	BE392080 601307712
41	557	96.3	734	7	BE743613	BE743613 601573751
42	583	96.3	734	7	BI819626	BI819626 603034162
43	770	96.2	733	4	BG484664	BG484664 602505842
c 44	728	95.5	728	5	BE271401	BE271401 601140427
45	511	95.3	726	1	CD599640	CD599640 EST16036
46	634	95.3	726	1	BG617592	BG617592 602615084
47	726	94.8	722	1	BG567519	BG567519 602586258
48	493	94.6	721	7	CD683966	CD683966 EST486 hu
49	560	94.5	720	4	BF038352	BF038352 601458220
c 50	447	94.1	717	5	BM708788	BM708788 UI-E-CI1-
c 51	584	94.0	716	1	AI638119	AI638119 ts9e012.x
c 52	582	93.6	713	5	AI885550	AI885550 wm25e06.x
53	394	93.2	713	2	AY417769	AY417769 Pan trogl
c 54	587	92.7	710	2	AI862145	AI862145 tw37d02.x
55	464	92.3	706	1	AA132992	AA132992 zo22a01.r
c 56	738	92.3	703	4	BF036845	BF036845 601460315
57	528	91.3	703	5	AA683520	AA683520 zf34b11.s
c 58	560	90.7	696	1	AW238407	AW238407 xp25b07.x
59	464	90.7	691.5	1	BE184455	BE184455 CM2-HT067
c 60	568	90.2	691	5	AI743345	AI743345 wg91h11.x
61	548	90.0	687	2	CD723292	CD723292 oj20d05.y
62	991	89.9	687	2	BE788273	BE788273 601481952
63	697	89.9	686	1	CO579827	CO579827 ILLUMIGEN
64	439	89.9	685	1	AA316675	AA316675 EST188351
c 65	469	89.8	685	4	BM708085	BM708085 UI-E-CI1-
c 66	553	88.7	685	7	BU731938	BU731938 UI-E-CI1-
c 67	561	88.4	684.5	1	AI742512	AI742512 wg43g10.x
c 68	537	88.2	680	1	AW081599	AW081599 xc32a03.x
69	548	88.2	676	5	BU730356	BU730356 UI-E-CI1-
c 70	451	88.2	673.5	1	BF896911	BF896911 CM2-MT015
71	505	88.2	672	5	BE612943	BE612943 601451888
72	519	87.5	672	1	BE612494	BE612494 601451888
73	584	87.1	672	4	AI564623	AI564623 tp98h03.x
c 74	739	87.1	672	7	BG537787	BG537787 60256614l
c 75	575	87.0	667	1	AA587140	AA587140 nn70006.s
76	456	86.9	664	1	R71834	R71834 yj86e07.r1
77	886	86.9	664	5	BG547643	BG547643 60257548
78	572	86.2	663	1	AI222907	AI222907 qm29e05.x
c 79	663	86.0	662.5	5	AA564454	AA564454 ng50e05.8
c 80	572	86.9	662	4	BU731488	BU731488 UI-E-CI1-
c 81	550	86.2	657	5	AA551908	AA551908 nk13b09.s
c 82	566	86.0	655	1	AA541595	AA541595 n189g08.8

```
c  83   654   85.8   BM694113   4   BM694113 UI-E-CI1-
c  84   652   85.6   AW513394   2   AW513394 xo44a01.x
c  85   650   85.3   BQ012955   2   BQ012955 UI-1-BC1p
c  86   648   85.0   BE933366   2   BE933366 RC0-HT088
c  87   643   84.4   AI580185   1   AI580185 tk17h06.x
c  88   643   84.4   AI858387   1   AI858387 w137e07.x
c  89   640   84.0   CD708989   6   CD708989 EST25516
c  90   640   84.0   AA863166   1   AA863166 og91d09.s
c  91   635   83.3   AA938592   1   AA938592 oo96d02.s
c  92   627   82.3   AA594366   1   AA594366 n197c05.s
c  93   624   81.9   BG678551   4   BG678551 602624231
c  94   623   81.8   BE740735   2   BE740735 601593520
c  95   622   81.6   BF032535   2   BF032535 601453382
c  96   619   81.2   AI148713   2   AI148713 qc65a01.x
c  97   617   81.0   BG194421   4   BG194421 RST13577
c  98   616   80.8   AI066534   1   AI066534 oo14e05.x
c  99   610   80.1   AI276644   1   AI276644 q164d02.x
c 100   607   79.7   BU674232   5   BU674232 UI-CF-DU0
```

ALIGNMENTS

```
RESULT 1
AA460433
LOCUS         AA460433         477 bp     mRNA      linear   EST 09-JUN-1997
DEFINITION    zx66h04.r1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
              IMAGE:796471 5' similar to gb:X04470 ANTILEUKOPROTEINASE 1
              (HUMAN);, mRNA sequence.
ACCESSION     AA460433
VERSION       AA460433.1  GI:2185179
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1  (bases 1 to 477)
AUTHORS       Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
              Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
              Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
              White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE         WashU-Merck EST Project 1997
JOURNAL       Unpublished (1997)
COMMENT       Contact: Wilson RK
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@watson.wustl.edu
              This clone is available royalty-free through LLNL ; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              Seq primer: -28m13 rev2 ET from Amersham
              High quality sequence stop: 416.
FEATURES             Location/Qualifiers
     source          1..477
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:6040683"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:796471"
                     /dev_stage="8-9 weeks"
                     /lab_host="DH10B"
                     /clone_lib="Soares_total_fetus_Nb2HF8_9w"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I, Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from pooled 8-9 week
                     (total) fetus material with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGAGCGGCCGCCTAATTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo. "
ORIGIN
```

```
Alignment Scores:
Pred. No.:                                 Length:             477
Score:                    2.82e-60         Matches:            132
Percent Similarity:       762.0            Conservative:         0
Best Local Similarity:    100.00%          Mismatches:           0
Query Match:              100.00%          Indels:               0
DB:                       100.00%          Gaps:                 0
                          1

US-10-613-105-2 (1-132) x AA460433 (1-477)

Qy    1 MetLysSerSerGlyLeuProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCT  61

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysSer  40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 121

Qy   41 AlaGlnCysLeuArgTyrLysLysProGlyLysCysGlnSerAspTrpGlnCysProGlyLys  60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 181

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  182 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 241

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  242 CCAACAAGGAGGAAGCCTGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 301

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  302 CCCCCAAATTTCTGTGAGATGGATGGGCAAGGCTGCAAGCGTGACTTGAAGTGTTGCATGGGC 361

Qy  121 MetCysGlyLysSerCysValSerProVallysAla 132
          |||||||||||||||||||||||||||||||||||
Db  362 ATGTGTGGAAATCTGCGGTTTCCCCGTGAAAGCT 397
```

```
RESULT 2
BI490742/c
LOCUS         BI490742                   534 bp    mRNA       linear   EST 28-AUG-2001
DEFINITION    603032084T1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173293 3',
              mRNA sequence.
ACCESSION     BI490742
VERSION       BI490742.1  GI:15329970
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1  (bases 1 to 534)
AUTHORS       NIH-MGC http://mgc.nci.nih.gov/.
TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL       Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM1431  row: g  column: 22
              High quality sequence start: 25
              High quality sequence stop: 534.
FEATURES             Location/Qualifiers
     source          1..534
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5173293"
```

```
                                      /lab_host="DH10B"
                                      /clone_lib="NIH_MGC 115"
                                      /note="Organ: pooled brain, lung, testis; Vector:
                                      pCMV-SPORT6; Site 1: NotI; Site_2: EcoRV (destroyed); RNA
                                      source anonymous pool of 6 male brains, age range 23-27; 1
                                      male lung, age 27; and 1 male testis, age 69. Library is
                                      oligo-dT primed and directionally cloned (EcoRV site is
                                      destroyed upon cloning). Average insert size is 1.8 kb,
                                      insert size range 1-3 kb. Library is normalized and
                                      enriched for full-length clones and was constructed by C.
                                      Gruber (Invitrogen). Research Genetics tracking code
                                      021. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.:            3.27e-60              Length:            534
Score:                762.00               Matches:           132
Percent Similarity:   100.00%              Conservative:      0
Best Local Similarity: 100.00%             Mismatches:        0
Query Match:          100.00%              Indels:            0
DB:                                        Gaps:              0

US-10-613-105-2 (1-132) x BI490742 (1-534)

Qy      1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    525 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCT 466

Qy     21 TrpAlaValGluGlySerGlyLysSerPheLysSerGlyAlaGlyValCysProProLysSer 40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    465 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 406

Qy     41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    405 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 346

Qy     61 LysArgCysCysProAspThrCysGlyIleLeuLysCysLysArgAspLeuLysCysProGly 80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    345 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 286

Qy     81 ProThrArgArgLysProGlyLysCysSerProValSerProValLysAla 132
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    285 CCAACAAGGAGGAAGGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 226

Qy    101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    225 CCCCCCAATTTCTGTGAGATGGACGGCCAGTGCAAGCGAGATCTGAAGTGTTGCATGGGC 166

Qy    121 MetCysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||||||||||||
Db    165 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 130


RESULT 3
BM769596
LOCUS       BM769596                 535 bp    mRNA      linear     EST 04-MAR-2002
DEFINITION  K-EST0052866 S14K402 Homo sapiens cDNA clone S14K402-16-B10 5',
            mRNA sequence.
ACCESSION   BM769596
VERSION     BM769596.1  GI:19009211
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 535)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
```

```
                                      52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
                                      Tel: +82-42-860-4470
                                      Fax: +82-42-860-4409
                                      Email: yongsung@mail.kribb.re.kr
                                      Plate: 16  row: B  column: 10
                                      High quality sequence stop: 535.

FEATURES        Location/Qualifiers
     source     1..535
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="S14K402-16-B10"
                /cell_line="K402"
                /clone_lib="S14K402"
                /note="Organ: Stomach; Vector: pT218RP1; Site 1: EcoRI;
                Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                bacterial alkaline phosphatase (BAP) and then decapped
                with tobacco acid pyrophosphatase (TAP). The decapped
                intact mRNA was ligated with DNA-RNA linker including EcoR
                I site by treatment of T4 RNA ligase and the first strand
                cDNA was synthesized from oligo dT-selected mRNA by
                priming with dT-tailed vector. The dT-tailed vector was
                adjusted to have about 60nt. The cDNA vector was
                circularized with E. coli DNA ligase after digestion of
                EcoRI which site is also included in vector. An RNA strand
                converted to a DNA strand by Okayama-Berg method. The
                obtained cDNA vectors were used for transformation of
                competent cells E. coli Top10F' by electroporation method.
                The cDNA libraries constructed by this method are
                full-length enriched cDNA library."

ORIGIN

Alignment Scores:
Pred. No.:            3.28e-60              Length:            535
Score:                762.00               Matches:           132
Percent Similarity:   100.00%              Conservative:      0
Best Local Similarity: 100.00%             Mismatches:        0
Query Match:          100.00%              Indels:            0
DB:                                        Gaps:              4

US-10-613-105-2 (1-132) x BM769596 (1-535)

Qy      1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     22 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCT 81

Qy     21 TrpAlaValGluGlySerGlyLysSerPheLysSerGlyAlaGlyValCysProProLysSer 40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     82 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 141

Qy     41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    142 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 201

Qy     61 LysArgCysCysProAspThrCysGlyIleLeuLysCysLysArgAspLeuLysCysProAsn 80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    202 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 261

Qy     81 ProThrArgArgLysProGlyLysCysSerProValThrProValGlyLeuMetLeuAsn 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    262 CCAACAAGGAGGAAGGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 321

Qy    101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    322 CCCCCCAATTTCTGTGAGATGGACGGCCAGTGCAAGCGAGATCTGAAGTGTTGCATGGGC 381

Qy    121 MetCysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||||||||||||
Db    382 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 417


RESULT 4
CD690842
```

LOCUS   CD690842   543 bp   mRNA   linear   EST 25-JUN-2003
DEFINITION   EST7365 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD690842
VERSION   CD690842.1   GI:32211980
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 543)
  AUTHORS   Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
          Zeng,Y.-X.
  TITLE   Transcriptional Gene Expression Profile of Human Nasopharynx
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Yixin Zeng
          Cancer Center
          Sun Yat-sen University
          651 DongFeng Road East, GuangZhou 510060, China
          Tel: 86-1380-9770-743
          Fax: 86-20-8775-4506
          Email: yxzeng@gzsums.edu.cn.
FEATURES   Location/Qualifiers
  source   1..543
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /tissue_type="normal nasopharynx"
          /clone_lib="human nasopharynx"
          /note="ESTs generated from a normal nasopharynx cDNA
          library from southern Chinese"
ORIGIN

Alignment Scores:
Pred. No.:   3.34e-60   Length: 543
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 6

US-10-613-105-2 (1-132) x CD690842 (1-543)

```
Qy    1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaAlaLeuGlyThrLeuAlaPro 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   53 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTGGCAGCTCTGGGAACTCTGGCACCT 112

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysSerProProLysLysSer 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  113 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCGTGTCCTCCTAAGAAATCT 172

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  173 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 232

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  233 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 292

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  293 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 352

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  353 CCCCCCAATTTCTGTGAGATGGATGGGCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 412

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||||||||||
Db  413 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 448
```

RESULT 5
CD707973
LOCUS   CD707973   554 bp   mRNA   linear   EST 25-JUN-2003

DEFINITION   EST24500 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD707973
VERSION   CD707973.1   GI:32238603
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 554)
  AUTHORS   Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
          Zeng,Y.-X.
  TITLE   Transcriptional Gene Expression Profile of Human Nasopharynx
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Yixin Zeng
          Cancer Center
          Sun Yat-sen University
          651 DongFeng Road East, GuangZhou 510060, China
          Tel: 86-1380-9770-743
          Fax: 86-20-8775-4506
          Email: yxzeng@gzsums.edu.cn.
FEATURES   Location/Qualifiers
  source   1..554
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /tissue_type="normal nasopharynx"
          /clone_lib="human nasopharynx"
          /note="ESTs generated from a normal nasopharynx cDNA
          library from southern Chinese"
ORIGIN

Alignment Scores:
Pred. No.:   3.43e-60   Length: 554
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 6

US-10-613-105-2 (1-132) x CD707973 (1-554)

```
Qy    1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaAlaLeuGlyThrLeuAlaPro 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   44 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTGGCAGCTCTGGGAACTCTGGCACCT 103

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysSerProProLysLysSer 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  104 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCGTGTCCTCCTAAGAAATCT 163

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  164 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 223

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  224 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 283

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  284 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 343

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  344 CCCCCCAATTTCTGTGAGATGGATGGGCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 403

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||||||||||
Db  404 ATGTGTGGGAAATCCTGCGTATCCCCTGTGAAAGCT 439
```

RESULT 6
BX089446
LOCUS   BX089446   563 bp   mRNA   linear   EST 23-JAN-2003
DEFINITION   BX089446 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone

ACCESSION   IMAGp998O15845 ; IMAGE:366950, mRNA sequence.
VERSION   BX089446
            BX089446.1 GI:27823826
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 563)
  AUTHORS   Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
            Radelof,U., Schneider,D. and Korn,B.
  TITLE   Human UnigeneSet - RZPD3
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD; IMAGp998O15845.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Human UnigeneSet - RZPD3 (RZPDLIB No.972)
            http://www.rzpd.de/CloneCards/cgi-
            bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.
FEATURES   Location/Qualifiers
     source   1..563
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGp998O15845 ; IMAGE:366950"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal heart NbHH19W"
                     /note="Organ: heart; Vector: pT7T3 (Pharmacia) with a
                     modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGAGCGGCGCCATCTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library was constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
ORIGIN

Alignment Scores:
Pred. No.:   563
Score: 3.51e-60   Matches: 132
Percent Similarity: 762.00   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 5

US-10-613-105-2 (1-132) x BX089446 (1-563)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTGGCCCTGGGAACTCTGGCACCT 60

Qy   21 TrpAlaValGluGlyLysSerGlyLysSerPheLysGlyAlaGlyValCysProProLysLysSer 40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TGGGCTGTGGAAGGCTTCGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 120

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||

```
DB:                              7            Gaps:            0

US-10-613-105-2 (1-132) x CK430164 (1-583)

Qy       1  MetLysSerGlyLeuPheProPheLeuValLeuAlaAlaLeuGlyThrLeuAlaPro  20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4  ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGGCCGCTGCCCTGGGAACTCTGGCACCT  63

Qy      21  TrpAlaValGluGluLysGlyLysSerPheLysAlaGlyValCysProProLysGlySer  40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64  TGGGCTGTGGAAGGCTCTGGAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT  123

Qy      41  AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124  GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG  183

Qy      61  LysArgCysThrCysGlyTyrIleLysLysAspAspValAspThrProAsn  80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     184  AAGAGATGTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC  243

Qy      81  ProThrArgLysProGlyLysProProValThrTyrGlyGlnLysLeuMetLeuAsn  100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     244  CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACGTTATGGCCAATGTTTGATGCTTAAC  303

Qy     101  ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     304  CCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC  363

Qy     121  MetCysGlyLysSerCysValSerProValLysAla  132
            ||||||||||||||||||||||||||||||||
Db     364  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  399
```

```
RESULT 8
BM767014
LOCUS       BM767014                593 bp    mRNA      linear  EST 04-MAR-2002
DEFINITION  K-EST0049047 S14K402 Homo sapiens cDNA clone S14K402-2-D07 5', mRNA
            sequence.
ACCESSION   BM767014
VERSION     BM767014.1  GI:1909629
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 593)
AUTHORS     Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE       21C Frontier Korean EST Project 2001
JOURNAL     Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 2  row: D  column: 07
            High quality sequence stop: 593.
FEATURES             Location/Qualifiers
     source          1..593
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S14K402-2-D07"
                     /cell_line="K402"
                     /lab_host="Top10F'"
                     /clone_lib="S14K402"
                     /note="Organ: Stomach; Vector: pT218RP1; Site_1: EcoRI;
                     Site_2: NotI; Organ: Stomach; Vector: pT218RP1; Site_1: EcoRI,
                     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including EcoR
```

```
                     I site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
                     EcoRI which site is also included in vector. An RNA strand
                     converted to a DNA strand by Okayama-Berg method. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."

ORIGIN

Alignment Scores:
Pred. No.:                 3.7e-60           Length:               593
Score:                     762.00            Matches:              132
Percent Similarity:        100.00%           Conservative:         0
Best Local Similarity:     100.00%           Mismatches:           0
Query Match:               100.00%           Indels:               0
DB:                                          Gaps:                 0

US-10-613-105-2 (1-132) x BM767014 (1-593)

Qy       1  MetLysSerGlyLeuPheProPheLeuValLeuAlaAlaLeuGlyThrLeuAlaPro  20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22  ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGGCCGCTGCCCTGGGAACTCTGGCACCT  81

Qy      21  TrpAlaValGluGluLysGlyLysSerPheLysAlaGlyValCysProProLysGlySer  40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82  TGGGCTGTGGAAGGCTCTGGAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT  141

Qy      41  AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142  GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG  201

Qy      61  LysArgCysThrCysGlyTyrIleLysLysAspAspValAspThrProAsn  80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     202  AAGAGATGTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC  261

Qy      81  ProThrArgLysProGlyLysProProValThrTyrGlyGlnLysLeuMetLeuAsn  100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     262  CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACGTTATGGCCAATGTTTGATGCTTAAC  321

Qy     101  ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     322  CCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC  381

Qy     121  MetCysGlyLysSerCysValSerProValLysAla  132
            ||||||||||||||||||||||||||||||||
Db     382  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  417
```

```
RESULT 9
BM843274
LOCUS       BM843274                597 bp    mRNA      linear  EST 06-MAR-2002
DEFINITION  K-EST0121093 S14K402 Homo sapiens cDNA clone S14K402-31-E12 5',
            mRNA sequence.
ACCESSION   BM843274
VERSION     BM843274.1  GI:19199697
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 597)
AUTHORS     Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE       21C Frontier Korean EST Project 2001
JOURNAL     Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
```

```
                    Tel: +82-42-860-4470
                    Fax: +82-42-860-4409
                    Email: yongsung@mail.kribb.re.kr
                    Plate: 31  row: E  column: 12
                    High quality sequence stop: 597.
FEATURES            Location/Qualifiers
     source         1..597
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="S14K402-31-E12"
                    /cell_line="K402"
                    /lab_host="Top10F'"
                    /clone_lib="S14K402"
                    /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
                    Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                    bacterial alkaline phosphatase (BAP) and then decapped
                    with tabacco acid pyrophosphatase (TAP). The decapped
                    intact mRNA was ligated with DNA-RNA linker including EcoR
                    I site by treatment of T4 RNA ligase and the first strand
                    cDNA was synthesized from oligo dT-selected mRNA by
                    priming with dT-tailed vector. The dT-tailed vector was
                    adjusted to have about 60nt. The cDNA vector was
                    circularized with E. coli DNA ligase after digestion of
                    EcoRI which site is also included in vector. An RNA strand
                    converted to a DNA strand by Okayama-Berg method. The
                    obtained cDNA vectors were used for transformation of
                    competent cells E. coli Top10F' by electroporation method.
                    The cDNA libraries constructed by this method are
                    full-length enriched cDNA library."

ORIGIN
     Alignment Scores:
     Pred. No.:            3.79e-60                  Length:         597
     Score:               762.00                     Matches:        132
     Percent Similarity:  100.00%                    Conservative:   0
     Best Local Similarity: 100.00%                  Mismatches:     0
     Query Match:         100.00%                     Indels:         0
     DB:                                             Gaps:           4

US-10-613-105-2 (1-132) x BM843274 (1-597)

Qy    1 MetLysSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro  20
Db   25 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGGCTGGAACTCTGGCACCT  84

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLeuAlaGlyValCysProLysSer  40
Db   85 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 144

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
Db  145 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 204

Qy   61 LysArgCysLysProAspThrCysSerGlyIleLysCysLeuAspProValAspThrProAsn  80
Db  205 AAGAGATGTTGTCCTGACACTTGTGCCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 264

Qy   81 ProThrArgArgLysProGlyLysProValThrTyrGlyLysGlnLysMetLeuAsn 100
Db  265 CCAACAAGAGGAGAAGCCTGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 324

Qy  101 ProProAsnPheCysGluMetAspGlyGlyLysLysLysSerArgArgLeuLysLysCysMetGly 120
Db  325 CCCCCAAATTTCTGTGAGATGGATGGCCAGCAAGCTGACTTGAAGTGTTGCATGGGC 384

Qy  121 MetCysGlyLysLysSerCysValSerProValLysAla 132
Db  385 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 420


RESULT 10
CD000587
LOCUS      CD000587            599 bp    mRNA    linear    EST 01-MAY-2003
```

```
DEFINITION  AGENCOURT 13548070 NIH_MGC_186 Homo sapiens cDNA clone
            IMAGE:30323772 5', mRNA sequence.
ACCESSION   CD000587
VERSION     CD000587.1  GI:30295106
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 599)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDCM126  row: o  column: 13
            High quality sequence stop: 556.
FEATURES            Location/Qualifiers
     source         1..599
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:30323772"
                    /lab_host="DH10B (T1 phage-resistant)"
                    /clone_lib="NIH_MGC_186"
                    /note="Organ: Pooled-Skin; Vector: pDNR-LIB; Site 1: SfiI
                    (ggccattatggcc); Site 2: SfiI (ggccgcctcggcc); Library is
                    oligo-dT primed and directionally cloned. cDNA was
                    prepared from a pooled samples of tissues from Skin,
                    meninges, duramatter, pia matter and choroid plexus.  5'
                    and 3' adaptors were used in cloning as follows: 5'
                    adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor
                    sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                    (where B = A, C, or G and N = A, C, G, or T). Average
                    insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies
                    contained inserts by PCR. This library was enriched for
                    full-length clones and was constructed by Clontech
                    Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
                    Library"

ORIGIN
     Alignment Scores:
     Pred. No.:            3.8e-60                   Length:         599
     Score:               762.00                     Matches:        132
     Percent Similarity:  100.00%                    Conservative:   0
     Best Local Similarity: 100.00%                  Mismatches:     0
     Query Match:         100.00%                     Indels:         0
     DB:                                             Gaps:           6

US-10-613-105-2 (1-132) x CD000587 (1-599)

Qy    1 MetLysSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
Db   25 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTGGCTGGAACTCTGGCACCT  84

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProLysSer  40
Db   85 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 144

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
Db  145 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 204

Qy   61 LysArgCysLysProAspThrCysSerGlyIleLysCysLeuAspProValAspThrProAsn  80
Db  205 AAGAGATGTTGTCCTGACACTTGTGCCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 264
```

```
Qy   81  ProThrArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  265  CCAACAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGCCAATGTTTGATGCTTAAC 324

Qy  101  ProProAsnPheCysGluMetAspGlyLysArgAspLeuLysCysMetGly 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  325  CCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 384

Qy  121  MetCysGlyLysSerValSerProValLysAla 132
         |||||||||||||||||||||||||||||
Db  385  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 420
```

RESULT 11
```
LOCUS       CK429948     603 bp    mRNA    linear    EST 06-JAN-2004
DEFINITION  oj43c03.y1 Human lacrimal gland, unamplified: oj Homo sapiens cDNA
            clone oj43c03 5', mRNA sequence.
ACCESSION   CK429948
VERSION     CK429948.1  GI:40677549
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 603)
AUTHORS     Dickinson,D., Laurie,G. and Wistow,G.
TITLE       Expressed sequence tag analysis of human lacrimal gland
JOURNAL     Unpublished (2002)
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 43  row: G  column: 03
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1..603
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="oj43c03"
                     /tissue_type="lacrimal gland"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human lacrimal gland, unamplified: oj"
                     /note="Organ: Eye; Vector: pCMVsport6; RNA was extracted
                     from 2 human lacrimal glands. A directionally cloned cDNA
                     library in the pCMVSport6 vector(life Technologies) was
                     constructed at Bioserve Biotechnology (Laurel MD)
                     essentially following the protocols of the SuperScript
                     Plasmid System full details of which are contained in the
                     manufacturer's instruction manual
                     (http://www.lifetech.com/). First strand synthesis was
                     carried out using a Not I primer-adapter
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCC(T)15-3']. EST analysis
                     was performed on the unamplified library at the NIH
                     Intramural Sequencing Center (NISC)."
ORIGIN
```

```
Alignment Scores:
Pred. No.:
Score:                  762.00       Length:            603
                        3.84e-60     Matches:           132
Percent Similarity:     100.00%      Conservative:      0
Best Local Similarity:  100.00%      Mismatches:        0
Query Match:            100.00%      Indels:            0
DB:                                  Gaps:              7
```

US-10-613-105-2 (1-132) x CK429948 (1-603)

```
Qy    1  MetLysSerSerGlyLeuPheProPheValLeuLeuAlaLeuAlaLeuGlyThrLeuAlaPro 20

Db    3  ATGAAGTCCAGCGGCCTCTTCCCCGTTCCCCTCGTGCTGGCTCTGGCTGCCCTGGGAACTCTCGGACCT 62

Qy   21  TrpAlaValGluGlyLysSerGlyLysSerPheLysAlaGlyLysValCysProLysLysSer 40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63  TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 122

Qy   41  AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123  GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTGACTGTCCAGTGTTCAGGGAAG 182

Qy   61  LysArgCysProAspThrCysGlyGlyIleLeuCysLeuAspAspProValAspThrProAsn 80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  183  AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 242

Qy   81  ProThrArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  243  CCAACAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGCCAATGTTTGATGCTTAAC 302

Qy  101  ProProAsnPheCysGluMetAspGlyLysArgAspLeuLysCysMetGly 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  303  CCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 362

Qy  121  MetCysGlyLysSerValSerProValLysAla 132
         |||||||||||||||||||||||||||||
Db  363  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 398
```

RESULT 12
```
LOCUS       BM842052     606 bp    mRNA    linear    EST 06-MAR-2002
DEFINITION  K-EST0119384 S12SNU216 Homo sapiens cDNA clone S12SNU216-57-G05 5',
            mRNA sequence.
ACCESSION   BM842052
VERSION     BM842052.1  GI:19198461
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 606)
AUTHORS     Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE       21C Frontier Korean EST Project 2001
JOURNAL     Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 57  row: G  column: 05
            High quality sequence stop: 606.
FEATURES             Location/Qualifiers
     source          1..606
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S12SNU216-57-G05"
                     /sex="F"
                     /tissue_type="Lymph node"
                     /cell_type="Epithelial"
                     /cell_line="SNU-216"
                     /lab_host="Top10F'"
                     /clone_lib="S12SNU216"
                     /note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including EcoR
                     I site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized from oligo dT-selected mRNA by
```

priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA libraries were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Alignment Scores:
Pred. No.:   3.86e-60
Score: 762.00   Length: 606
Percent Similarity: 100.00%   Matches: 132
Best Local Similarity: 100.00%   Conservative: 0
Query Match: 100.00%   Mismatches: 0
DB: Indels: 0
                                       Gaps: 4

US-10-613-105-2 (1-132) x BM842052 (1-606)

Qy   1 MetLysSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGTTGGCTCTGGGAACTCTGGCACCT 81

Qy   21 TrpAlaValGluGluSerGlyLysSerPheLysAlaGlyValCysProLysLysSer 40
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 141

Qy   41 AlaGlnCysLeuArgTyrLysProGluCysTrpSerAspTrpGlnCysProGlyLys 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   142 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 201

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   202 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 261

Qy   81 ProThrArgArgLysProGlyLysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   262 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 321

Qy   101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuCysCysMetGly 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   322 CCCCCAAATTTCTGTGAGATGGATGGCCAGTGCAAGAAGCGTGACTTGAAGTGTTGCATGGGC 381

Qy   121 MetCysGlyLysSerCysValSerProValLysAla 132
           |||||||||||||||||||||||||||||||||
Db   382 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 417

RESULT 13
BM976429/c
LOCUS   BM976429   612 bp   mRNA   linear   EST 21-FEB-2003
DEFINITION UI-CF-EN1-acz-n-15-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
            UI-CF-EN1-acz-n-15-0-UI 3', mRNA sequence.
ACCESSION   BM976429
VERSION   BM976429.1 GI:19594023
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 612)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE   Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   MEDLINE   97044477
   PUBMED   8889548
COMMENT   Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866

Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com) or from Open Biosystems
            (www.openbiosystems.com).
            Seq primer: M13 FORWARD
            POLYA=Yes.

FEATURES   Location/Qualifiers
     source   1..612
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-CF-EN1-acz-n-15-0-UI"
                     /tissue_type="Primary Lung Cystic Fibrosis Epithelial
                     Cells"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-CF-EN1"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-CF-EN1 is a normalized cDNA library containing the
                     following tissue(s): Primary Lung Cystic Fibrosis
                     Epithelial Cells. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pT7T3-Pac vector. The
                     oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CTGCTCAGGT.
                     TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
                     6hr to LPS 24h
                     TAG_LIB=UI-CF-EN1
                     TAG_SEQ=CTGCTCAGGT"

ORIGIN

Alignment Scores:
Pred. No.:   3.91e-60
Score: 762.00   Length: 612
Percent Similarity: 100.00%   Matches: 132
Best Local Similarity: 100.00%   Conservative: 0
Query Match: 100.00%   Mismatches: 0
DB: Indels: 0
                                       Gaps: 5

US-10-613-105-2 (1-132) x BM976429 (1-612)

Qy   1 MetLysSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   592 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGTTGGCTCTGGGAACTCTGGCACCT 533

Qy   21 TrpAlaValGluGluSerGlyLysSerPheLysAlaGlyValCysProLysLysSer 40
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   532 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 473

Qy   41 AlaGlnCysLeuArgTyrLysProGluCysTrpSerAspTrpGlnCysProGlyLys 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   472 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 413

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   412 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 353

Qy   81 ProThrArgArgLysProGlyLysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   352 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 293

Qy   101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuCysCysMetGly 120

Db   191 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCAAAC 250
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   292 CCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 233

Qy   81 ProThrArgLysGlyLysProValThrTyrGlnCysLeuMetLeuAsn 100
             ||| ||||||||||||||||||||||||||||||||||||||||||||
Qy   121 MetCysGlyLysSerCysValSerProValLysAla 132

Db   251 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGCTTATGGCCAATGTTGATGCTTAAC 310
             ||||||||||||||||||||||||||||||||||||||||
Db   232 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 197

Qy   101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysAspAspLeuLysCysCysMetGly 120

Db   311 CCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 370

Qy   121 MetCysGlyLysSerCysValSerProValLysAla 132

Db   371 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 406

RESULT 15
CD103997
LOCUS   CD103997   760 bp   mRNA   linear   EST 15-MAY-2003
DEFINITION   AGENCOURT 13980945 NIH_MGC_186 Homo sapiens cDNA clone
             IMAGE:30371083 5', mRNA sequence.
ACCESSION   CD103997
VERSION   CD103997.1 GI:30757171
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 760)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: NDCM161 row: b column: 20
             High quality sequence stop: 458.
FEATURES   Location/Qualifiers
     source   1..760
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30371083"
                     /clone_lib="NIH_MGC_186"
                     /note="Organ: Pooled-Skin; Vector: pDNR-LIB; Site_1: SfiI
                     (ggccattaggcc); Site_2: SfiI (ggccgcctcggcc); Library is
                     oligo-dT primed and directionally cloned. cDNA was
                     prepared from a pooled samples of tissues from Skin,
                     meninges, duramatter, pia matter and choroid plexus. 5'
                     and 3' adaptors were used in cloning as follows: 5' adaptor
                     sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C, G, or T). Average
                     insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies
                     contained inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
                     Library"

ORIGIN

Alignment Scores:
Pred. No.:   Length: 760
Score: 5.2e-60   Matches: 132
Percent Similarity: 762.00   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 6

Db   292 CCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 233
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   121 MetCysGlyLysSerCysValSerProValLysAla 132
             |||||||||||||||||||||||||||||||||||
Db   232 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 197

RESULT 14
BQ232871
LOCUS   BQ232871   626 bp   mRNA   linear   EST 02-MAY-2002
DEFINITION   AGENCOURT 7579635 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6057322
             5', mRNA sequence.
ACCESSION   BQ232871
VERSION   BQ232871.1 GI:2041427l
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 626)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLAM13320 row: j column: 11
             High quality sequence stop: 546.
FEATURES   Location/Qualifiers
     source   1..626
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6057322"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "

ORIGIN

Alignment Scores:
Pred. No.:   Length: 626
Score: 4.03e-60   Matches: 132
Percent Similarity: 762.00   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 5   Gaps: 5

US-10-613-105-2 (1-132) x BQ232871 (1-626)

Qy   1 MetLysSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11 ATGAAGTCCAAGGCGCCTCTTCCCCTTCCCCTGTCGTGCTTGCCCTGGCAACTCTGCCACCT 70

Qy   21 TrpAlaValGluGluSerGlyLysSerGlyLysSerPheLysAlaGlyValCysProProLysSer 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   71 TGGGCTGTGGAAGAGTCTGGAAAGTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 130

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnSerProGlyLys 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   131 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 190

Qy   61 LysArgCysGlyProAspThrCysGlyIleLysCysCysAspProValAspThrProAsn 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
US-10-613-105-2 (1-132) x CID103997 (1-760)

Qy     1 MetLysSerSerGlyLeuPheProPheValLeuLeuAlaLeuGlyThrLeuAlaPro  20
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAAGTCCAGCGGCCTCTTCCCCTTCGTGCTCCTGGGAACTCTGGCACCT      85

Qy    21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProLysSer     40
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    86 TGGGCTGTGGAAGGCTCGGAAAGTCCTTCAAAGCTGGAGTCTGCTCCTAAGAAATCT    145

Qy    41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   146 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG   205

Qy    61 LysArgCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn     80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   206 AAGAGATGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC    265

Qy    81 ProThrArgLysProGlyLysProValThrTyrGlyGlnCysLeuMetLeuAsn    100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   266 CCAACAAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC    325

Qy   101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysMetGly  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   326 CCCCCAAATTTCTGTGAGATGGACGGCCAAGCGTGCAAGCGTGACTTGAAGTGTCATGGGC    385

Qy   121 MetCysGlyLysSerCysValSerProValLysAla  132
         |||||||||||||||||||||||||||||||||||||
Db   386 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT    421
```

```
RESULT 16
CD000527
LOCUS       CD000527      796 bp    mRNA       linear   EST 01-MAY-2003
DEFINITION  AGENCOURT_13651909 NIH_MGC_186 Homo sapiens cDNA clone
            IMAGE:30321423 5', mRNA sequence.
ACCESSION   CD000527
VERSION     CD000527.1  GI:30295046
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 796)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
            cDNA Library Preparation: CLONETECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDCM120 row: m  column: 16
            High quality sequence stop: 485.
FEATURES             Location/Qualifiers
     source          1..796
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30321423"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_186"
                     /note="Organ: Pooled-Skin; Vector: pDNR-LIB; Site 1: SfiI
                     (ggccattaggcc); Site 2: SfiI (ggccgcctcggcc); Library is
                     oligo-dT primed and directionally cloned. cDNA was
                     prepared from a pooled samples of tissues from Skin,
                     meninges, duramatter, pia matter and choroid plexus.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
```

(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library"

```
ORIGIN

Alignment Scores:
Pred. No.:                5.52e-60    Length:      796
Score:                    762.00      Matches:     132
Percent Similarity:       100.00%     Conservative:  0
Best Local Similarity:    100.00%     Mismatches:    0
Query Match:              100.00%     Indels:        0
DB:                                   Gaps:          6

US-10-613-105-2 (1-132) x CD000527 (1-796)

Qy     1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    26 ATGAAGTCCAGCGGCCCTCTTCCCCTTCGTGCTCCTGGGAACTCTGGCACCT      85

Qy    21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProLysSer     40
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    86 TGGGCTGTGGAAGGCTCGGAAAGTCCTTCAAAGCTGGAGTCTGCTCCTAAGAAATCT    145

Qy    41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   146 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG   205

Qy    61 LysArgCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn     80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   206 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC  265

Qy    81 ProThrArgLysProGlyLysProValThrTyrGlyGlnCysLeuMetLeuAsn    100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   266 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC  325

Qy   101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysMetGly  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   326 CCCCCCAATTTCTGTGAGATGGATGGCCAAGCGTGCAAGCGTGACTTGAAGTGTCATGGGC  385

Qy   121 MetCysGlyLysSerCysValSerProValLysAla  132
         |||||||||||||||||||||||||||||||||||||
Db   386 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT    421
```

```
RESULT 17
BG533465
LOCUS       BG533465      874 bp    mRNA       linear   EST 03-APR-2001
DEFINITION  601860869F2 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4072049 5',
            mRNA sequence.
ACCESSION   BG533465
VERSION     BG533465.1  GI:13525005
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 874)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONETECH Laboratories, Inc.
            cDNA Library Preparation: CLONETECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM918  row:j  column: 18
            High quality sequence stop: 594.
```

FEATURES
  source   Location/Qualifiers
    1..874
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:4072049"
    /lab_host="DH10B (T1 phage-resistant)"
    /clone_lib="NIH_MGC_76"
    /note="Organ: liver; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccattatggcc); Site_2: SfiI (ggccgcctcggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.85 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.:   6.24e-60   Length: 874
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 4

US-10-613-105-2 (1-132) x BG553465 (1-874)

Qy   1 MetLysSerSerGlyLysPhePheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
Db   23 ATGAAGTCCAGCGGCCGCCTCTTCCCCTTCCTGGTGCTCGTGCTTGCCCTGGGAACTCTGGCACCT 82

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer 40
Db   83 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 142

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
Db 143 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 202

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
Db 203 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTTGACCCTGTTGACACCCCAAAC 262

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
Db 263 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAAATGTTTGATGCTTAAC 322

Qy 101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
Db 323 CCCCCAAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 382

Qy 121 MetCysGlyLysSerCysValSerProValLysAla 132
Db 383 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 418

RESULT 18
AY417768
LOCUS   AY417768   394 bp   DNA   linear   GSS 17-DEC-2003
DEFINITION   Homo sapiens SLPI gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION   AY417768
VERSION   AY417768.1   GI:39773728
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 394)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

           Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL   Science 302 (5652), 1960-1963 (2003)
PUBMED   14671302
REFERENCE   2 (bases 1 to 394)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE   Direct Submission
JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES   Location/Qualifiers
  source   1..394
           /organism="Homo sapiens"
           /mol_type="genomic DNA"
           /db_xref="taxon:9606"
  gene   <1..>394
           /gene="SLPI"
           /locus_tag="HCM6334"
ORIGIN

Alignment Scores:
Pred. No.:   5.12e-60   Length: 394
Score: 758.00   Matches: 131
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 99.48%   Indels: 0
DB: Gaps: 0

US-10-613-105-2 (1-132) x AY417768 (1-394)

Qy   1 MetLysSerSerGlyLysPhePheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
Db   1 ATGAAGTCCAGCGGCCGCCTCTTCCCCTTCCTGGTGCTCGTGCTTGCCCTGGGAACTCTGGCACCT 60

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer 40
Db   61 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 120

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
Db 121 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 180

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
Db 181 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTTGACCCTGTTGACACCCCAAAC 240

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
Db 241 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAAATGTTTGATGCTTAAC 300

Qy 101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
Db 301 CCCCCAAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 360

Qy 121 MetCysGlyLysSerCysValSerProValLys 131
Db 361 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAA 393

RESULT 19
CF529326/c
LOCUS   CF529326   597 bp   mRNA   linear   EST 11-SEP-2003
DEFINITION   UI-1-BC1p-atb-c-09-0-UI.s1 NCI_CGAP_P13 Homo sapiens cDNA clone UI-1-BC1p-atb-c-09-0-UI 3', mRNA sequence.
ACCESSION   CF529326
VERSION   CF529326.1   GI:34579021
KEYWORDS   EST.
SOURCE   Homo sapiens (human)

Left record

```
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 597)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Steven Brown
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES    Location/Qualifiers
   source   1..597
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="UI-1-BC1p-atb-c-09-0-UI"
            /tissue_type="Placenta"
            /dev_stage="8-9 weeks"
            /lab_host="DH10B (Life Technologies)"
            /clone_lib="NCI_CGAP P13"
            /note="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with
            a modified polylinker; Site 1: EcoR I; Site 2: Not I;
            NCI_CGAP_P13 is a subtracted cDNA library constructed
            according to Bonaldo, Lennon and Soares, Genome Research,
            6:791-806, 1996. First strand cDNA synthesis was primed
            with an oligo-dT primer containing a Not I site. Double
            stranded cDNA was ligated to an EcoR I adaptor, digested
            with Not I, and cloned directionally into pT7T3-Pac
            vector. The oligonucleotide used to prime the synthesis of
            first-strand cDNA contains a library tag sequence that is
            located between the Not I site and the (dT)18 tail. The
            sequence tags for this library are GA, AGGAA. For
            additional information, contact: Bento Soares,
            bento-soares@uiowa.edu
            TAG_TISSUE=placenta human 8 week
            TAG_LIB=UI-1-BC1p
            TAG_SEQ=GA"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              8.84e-60       Length:                 597
Score:                  758.00         Matches:                131
Percent Similarity:     100.00%        Conservative:           1
Best Local Similarity:  99.24%         Mismatches:             0
Query Match:            99.48%         Indels:                 0
DB:                     7              Gaps:                   0

US-10-613-105-2 (1-132) x CF529326 (1-597)
```

```
Qy    1  MetLysSerSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro  20
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  591  GTGAAGTCCAGCGGCCGCTCTTCCCCTTCCTGGTCGTTGCCCTGGGAACTCTGGCACCT  532

Qy   21  TrpAlaValGluGluSerGlyGlyLysSerPheLysSerPheLysAlaGlyTyrValCysProProLysLysSer  40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  531  TGGGCTGTGGAAGGCTTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT  472

Qy   41  AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerSerAspTrpGlnCysProGlyLys  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  471  GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG  412

Qy   61  LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  411  AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC  352
```

Right record

```
Qy   81  ProThrArgArgLysProGlyLysCysProValThrTyrGlnGlnCysLeuMetLeuAsn 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  351  CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTGATGCTTAAC  292

Qy  101  ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  291  CCCCCAATTTCTGTGAGATGATGGCCAGTGCAAGCGTGACTTGAGTGTTGCATGGGC  232

Qy  121  MetCysGlyLysSerCysValSerProValLysAla 132
         ||||||||||||||||||||||||||||||||||||
Db  231  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  196
```

```
RESULT 20
BU753561/c
LOCUS       BU753561        598 bp    mRNA    linear    EST 10-OCT-2002
DEFINITION  UI-1-BC1-ajl-e-04-0-UI.s1 NCI_CGAP P12 Homo sapiens cDNA clone
            UI-1-BC1-ajl-e-04-0-UI 3', mRNA sequence.
ACCESSION   BU753561
VERSION     BU753561.1  GI:23711685
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 598)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Steven Brown
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES    Location/Qualifiers
   source   1..598
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="UI-1-BC1-ajl-e-04-0-UI"
            /tissue_type="placenta"
            /dev_stage="8-9 weeks"
            /lab_host="DH10B (Life Technologies)"
            /clone_lib="NCI_CGAP P12"
            /note="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with
            a modified polylinker; Site 1: EcoR I; Site 2: Not I;
            NCI_CGAP_P12 is a subtracted cDNA library constructed
            according to Bonaldo, Lennon and Soares, Genome Research,
            6:791-806, 1996. First strand cDNA synthesis was primed
            with an oligo-dT primer containing a Not I site. Double
            stranded cDNA was ligated to an EcoR I adaptor, digested
            with Not I, and cloned directionally into pT7T3-Pac
            vector. The oligonucleotide used to prime the synthesis of
            first-strand cDNA contains a library tag sequence that is
            located between the Not I site and the (dT)18 tail. The
            sequence tags for this library are GA, AGGAA. For
            additional information, contact: Bento Soares,
            bento-soares@uiowa.edu
            TAG_TISSUE=placenta human 8 week
            TAG_LIB=UI-1-BC1
            TAG_SEQ=GA"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              8.86e-60       Length:                 598
Score:                  758.00         Matches:                131
Percent Similarity:     100.00%        Conservative:           1
Best Local Similarity:  99.24%         Mismatches:             0
```

```
Query Match:       99.48%      Indels:       0
Db:                5           Gaps:         0

US-10-613-105-2 (1-132) x BU753561 (1-599)

Qy       1 MetLysSerSerGlyLeuPheProPheValLeuLeuAlaLeuGlyThrLeuAlaPro 20
           :::|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     591 GTGAAGTCCAGCGGCCTCTTCCCTGTGCTTGCCCTGGGAACTCTGGCCACCT 532

Qy      21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyLysValCysProProLysLysSer 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     531 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 472

Qy      41 AlaGlnCysLeuAlaArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     471 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 412

Qy      61 LysArgCysProAspThrCysGlyIleLysCysLeuAspValAspThrProAsn 80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     411 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATGTTGACACCCCAAAC 352

Qy      81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     351 CCAACAAGGAGGAAGCCTGGAAGTGCCCAGTGACTTATGCCAATGTTTGATGCTTAAC 292

Qy     101 ProProAsnPheCysGluMetAspGlyLysGlnCysLysArgAspLeuLysCysCysMetGly 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     291 CCCCCAATTTCTGTGAGATGGATGGCAAGCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 232

Qy     121 MetCysGlyLysSerCysValSerProValLysAla 132
           |||||||||||||||||||||||||||||||||||||
Db     231 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 196

RESULT 21
CB305723/c
LOCUS       CB305723                         599 bp    mRNA           linear   EST 04-MAR-2003
DEFINITION  UI-CF-EN1-aeg-o-21-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
            UI-CF-EN1-aeg-o-21-0-UI 3', mRNA sequence.
ACCESSION   CB305723
VERSION     CB305723.1  GI:2884623
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 599)
AUTHORS     Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     97044477
PUBMED      8889548
COMMENT     Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com) or from Open Biosystems
            (www.openbiosystems.com).
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES            Location/Qualifiers
     source         1..599
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
```

```
                    /clone="UI-CF-EN1-aeg-o-21-0-UI"
                    /tissue_type="Primary Lung Cystic Fibrosis Epithelial
                    Cells"
                    /dev_stage="Adult"
                    /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                    /clone_lib="UI-CF-EN1"
                    /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                    modified polylinker; Site_1: EcoR I; Site_2: Not I;
                    UI-CF-EN1 is a normalized cDNA library containing the
                    following tissue(s): Primary Lung Cystic Fibrosis
                    Epithelial Cells. The library was constructed according to
                    Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                    1996. First strand cDNA synthesis was primed with an
                    oligo-dT primer containing a Not I site.  Double stranded
                    cDNA was ligated to an EcoR I adaptor, digested with Not
                    I, and cloned directionally into pT7T3-Pac vector. The
                    oligonucleotide used to prime the synthesis of
                    first-strand cDNA contains a library tag sequence that is
                    located between the Not I site and the (dT)18 tail. The
                    sequence tag for this library is CTGCTCAGGT.
                    TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
                    6hr to LPS 24h
                    TAG_LIB=UI-CF-EN1
                    TAG_SEQ=CTGCTCAGGT"

ORIGIN

Alignment Scores:
Pred. No.:            8.88e-60            Length:              599
Score:                758.00              Matches:             131
Percent Similarity:   100.00%             Conservative:        1
Best Local Similarity: 99.24%             Mismatches:          0
Query Match:          99.48%             Indels:              0
Db:                                       Gaps:                0

US-10-613-105-2 (1-132) x CB305723 (1-599)

Qy       1 MetLysSerSerGlyLeuPheProPheValLeuLeuAlaLeuGlyThrLeuAlaPro 20
           :::|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     592 GTGAAGTCCAGCGGCCTCTTCCCTGTGCTTGCCCTGGGAACTCTGGCCACCT 533

Qy      21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyLysValCysProProLysLysSer 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     532 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 473

Qy      41 AlaGlnCysLeuAlaArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     472 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 413

Qy      61 LysArgCysProAspThrCysGlyIleLysCysLeuAspValAspThrProAsn 80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     412 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATGTTGACACCCCAAAC 353

Qy      81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     352 CCAACAAGGAGGAAGCCTGGAAGTGCCCAGTGACTTATGCCAATGTTTGATGCTTAAC 293

Qy     101 ProProAsnPheCysGluMetAspGlyLysGlnCysLysArgAspLeuLysCysCysMetGly 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     292 CCCCCAATTTCTGTGAGATGGATGGCAAGCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 233

Qy     121 MetCysGlyLysSerCysValSerProValLysAla 132
           |||||||||||||||||||||||||||||||||||||
Db     232 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 197

RESULT 22
CF529317/c
LOCUS       CF529317                         607 bp    mRNA           linear   EST 11-SEP-2003
DEFINITION  UI-1-BC1p-atb-b-03-0-UI.s1 NCI_CGAP_P13 Homo sapiens cDNA clone
            UI-1-BC1p-atb-b-03-0-UI 3', mRNA sequence.
ACCESSION   CF529317
VERSION     CF529317.1  GI:34579012
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
```

```
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 607)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Steven Brown
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1..607
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-1-BCip-atb-b-03-0-UI"
                     /tissue_type="Placenta"
                     /dev_stage="8-9 weeks"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_P13"
                     /note="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with
                     a modified polylinker; Site 1: EcoR I; Site 2: Not I;
                     NCI_CGAP_P13 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are GA, AGGAA. For
                     additional information, contact: Bento Soares,
                     bento-soares@uiowa.edu
                     TAG_TISSUE=placenta human 8 week
                     TAG_LIB=UI-1-BCip
                     TAG_SEQ=GA"
ORIGIN

Alignment Scores:
Pred. No.:                9.03e-60          Length:              607
Score:                    758.00            Matches:             131
Percent Similarity:       100.00%           Conservative:        1
Best Local Similarity:    99.24%            Mismatches:          0
Query Match:              99.48%            Indels:              0
DB:                                         Gaps:                0

US-10-613-105-2 (1-132) x CF529317 (1-607)

Qy     1 MetLysSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro  20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   592 GTGAAGTCCAGCGGCCCTTCTCCCCTTCCCCTTGGTGTGCTTGCCCTGGGAACTCTGGCACCT 533

Qy    21 TrpAlaValGluGlyLysGlyLysSerPheLysAlaGlyValCysProLysLysSer  40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   532 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTGTGTCCTCCTAAGAAATCT 473

Qy    41 AlaGlnLeuArgTyrLysLysProGluCysGlnSerAspTrpProGlyGlyLys  60
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   472 GCCCAGTTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCCAGTGTCCAGGGAAG 413

Qy    61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   412 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGACCCCCAAAC 353
```

```
Qy    81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuLeuMetLeuAsn 100
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   352 CCACACAAGGAGGAGGCCTGGGAGTGCCCAGTGACTTATGGCCAATGTTGATGCTTAAC 293

Qy   101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly 120
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   292 CCCCCAATTTCTGTGAGATGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 233

Qy   121 MetCysGlyLysSerCysValSerProValLysAla 132
         ||||||||||||||||||||||||||
Db   232 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 197

RESULT 23
CD685198
CD685198
LOCUS       CD685198                623 bp    mRNA      linear      EST 25-JUN-2003
DEFINITION  EST1718 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD685198
VERSION     CD685198.1  GI:33200931
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 623)
AUTHORS     Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
            Zeng,Y.x.
TITLE       Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL     Unpublished (2003)
COMMENT     Contact: Yixin Zeng
            Cancer Center
            Sun Yat-sen University
            651 DongFeng Road East, GuangZhou 510060, China
            Tel: 86-1380-9770-743
            Fax: 86-20-8775-4506
            Email: yxzeng@gzsums.edu.cn.
FEATURES             Location/Qualifiers
     source          1..623
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="normal nasopharynx"
                     /clone_lib="human nasopharynx"
                     /note="ESTs generated from a normal nasopharynx cDNA
                     library from southern Chinese"
ORIGIN

Alignment Scores:
Pred. No.:                9.35e-60          Length:              623
Score:                    758.00            Matches:             131
Percent Similarity:       100.00%           Conservative:        1
Best Local Similarity:    99.24%            Mismatches:          0
Query Match:              99.48%            Indels:              0
DB:                                         Gaps:                0

US-10-613-105-2 (1-132) x CD685198 (1-623)

Qy     1 MetLysSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro  20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    44 ATGAAGTCCAGCGGCCCTTCTCCCCTTCCCCTTGGTGTGCTTGCCCTGGGAACTCTGGCACCT 103

Qy    21 TrpAlaValGluGlyLysGlyLysSerPheLysAlaGlyValCysProLysLysSer  40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   104 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTGTGTCCTCCTAAGAAATCT 163

Qy    41 AlaGlnLeuArgTyrLysLysProGluCysGlnSerAspTrpProGlyGlyLys  60
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   164 GCCCAGTTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCCAGTGTCCAGGGAAG 223

Qy    61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   224 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGACCCCCAAAC 283

Qy    81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuLeuMetLeuAsn 100
```

```
Db   284  CCAACAAGGAGGAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC  343
          ||||||||||||||||::||||||||||||||||||||||||| |||||||||||||||
Qy   101  ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysMetGly  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   344  CCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC  403
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121  MetCysGlyLysSerValSerProValLysAla  132
          |||||||||||||||||||||||||||||||||
Db   404  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  439
```

RESULT 24
CD689844
LOCUS   CD689844   524 bp   mRNA   linear   EST 25-JUN-2003
DEFINITION   EST6367 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD689844
VERSION   CD689844.1 GI:33210005
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 524)
AUTHORS   Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
             Zeng,Y.-X.
TITLE   Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL   Unpublished (2003)
COMMENT   Contact: Yixin Zeng
             Cancer Center
             Sun Yat-sen University
             651 DongFeng Road East, Guangzhou 510060, China
             Tel: 86-1380-9770-743
             Fax: 86-20-8775-4506
             Email: yxzeng@gzsums.edu.cn.
FEATURES   Location/Qualifiers
    source   1..524
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /tissue_type="normal nasopharynx"
             /clone_lib="human nasopharynx"
             /note="ESTs generated from a normal nasopharynx cDNA
             library from southern Chinese"

ORIGIN

Alignment Scores:
Pred. No.:   9.21e-60   Length: 524
Score: 757.00   Matches: 131
Percent Similarity: 99.24%   Conservative: 0
Best Local Similarity: 99.24%   Mismatches: 1
Query Match: 99.34%   Indels: 0
DB: Gaps: 0

US-10-613-105-2 (1-132) x CD689844 (1-524)

```
Qy    1  MetLysSerSerGlyLeuPheProPheLeuValLeuValLeuAlaLeuGlyThrLeuAlaPro  20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGGTGCTGGCCCTGGGAACTCTGGCACCT  122
Qy   21  TrpAlaValGluGlyLysPheLysSerGlyAlaGlyValCysProProLysGlySerSer  40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63  TGGGCTGTGGAAGGCTCGAAAGTCCTGGAAGTCCTCAAAGTCGGAGTCTGTCCTCCTAAGAAATCT  182
Qy   41  AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  183  GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTGCGAGTGCCAGTGTCCAGGGAAG  242
Qy   61  LysArgCysCysAspThrCysGlyIleLeuValCysLysLeuAspProValAspThrProAsn  80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  243  AAGAGATGTTGTCCTGACACTTGTGGGCATCCTGGTGTGCAAGCTTGACCCTGTTGACACCCCAAAC  302
Qy   81  ProThrArgArgLysProGlyGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn  100
         ||||||||||||
```

```
Db   303  CCAACAAGGAGGAGCCTGGGAAGTGCCCAGTGACTTATGGCCACTGTTTGATGCTTAAC  362
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   101  ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysMetGly  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   363  CCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC  422
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121  MetCysGlyLysSerValSerProValLysAla  132
          |||||||||||||||||||||||||||||||||
Db   423  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  458
```

RESULT 25
BQ436671
LOCUS   BQ436671   589 bp   mRNA   linear   EST 24-MAY-2002
DEFINITION   AGENCOURT_7776044 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6024344
             5', mRNA sequence.
ACCESSION   BQ436671
VERSION   BQ436671.1 GI:21175747
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 589)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLAM13234 row: 1 column: 09
             High quality sequence stop: 588.
FEATURES   Location/Qualifiers
    source   1..589
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:6024344"
             /tissue_type="large cell carcinoma"
             /lab_host="DH10B (phage-resistant)"
             /clone_lib="NIH_MGC_68"
             /note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
             Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
             Average insert size 1.8 kb. Library constructed by Life
             Technologies. "

ORIGIN

Alignment Scores:
Pred. No.:   1.07e-59   Length: 589
Score: 757.00   Matches: 131
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 99.34%   Indels: 0
DB: Gaps: 0

US-10-613-105-2 (1-132) x BQ436671 (1-589)

```
Qy    2  LysSerSerGlyLeuPheProPheLeuValLeuValLeuAlaLeuGlyThrLeuAlaProTrp  21
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3  AAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGGTGCTGGCCCTGGGAACTCTGGCACCTTGG  62
Qy   22  AlaValGluGlyLysPheLysSerGlyAlaGlyValCysProProLysGlyValCysSerSerAla  41
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63  GCTGTGGAAGGCTCGAAAGTCCTGGAAGTCGGAGTCTGTCCTCCTAAGAAATCTGCC  122
Qy   42  GlnCysLeuArgTyrLysLysProGluCysGluGlnSerAspTrpGlnCysProGlyLeuLys  61
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123  CAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTGCCAGTGGCAGTGTCCAGGGAAGAAG  182
```

```
Qy   62 ArgCysProAspThrCysGlyIleLeuCysLeuAspProValAspThrProAsnPro 81
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  183 AGATGTTGCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAACCCA 242

Qy   82 ThrArgLysProGlyLysCysProValThrTyrGlyIleGlnCysLeuMetLeuAsnPro 101
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  243 ACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAACCCC 302

Qy  102 ProAsnPheCysGluMetAspGlyGlnCysLysArgLeuLysCysGlyMet 121
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  303 CCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGCATG 362

Qy  122 CysGlyLysSerCysValSerProValLysAla 132
        |||||||||||||||||||||||||||||||||
Db  363 TGTGGGAAATCCTGCGTTCCCGTGAAAGCT 395
```

RESULT 26
BI253834
LOCUS   BI253834   608 bp   mRNA   linear   EST 17-JUL-2001
DEFINITION 602974607F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5113920 5',
            mRNA sequence.
ACCESSION   BI253834
VERSION   BI253834.1 GI:14805650
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 608)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Incyte Genomics, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11276 row: n column: 01
            High quality sequence stop: 567.
FEATURES   Location/Qualifiers
     source   1..608
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5113920"
                     /tissue_type="cervical carcinoma cell line"
                     /lab_host="DH10B"
                     /clone_lib="NIH MGC 12"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.4 kb. Library prepared by Life
                     Technologies."
ORIGIN

Alignment Scores:
Pred. No.:   1.12e-59   Length: 608
Score: 757.00   Matches: 131
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 99.34%   Indels: 0
DB: Gaps: 4

US-10-613-105-2 (1-132) x BI253834 (1-608)

Qy   2 LysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaProTrp 21
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3 AAGTCCAGCCGGCCTCTTCCCCCTTCCTGGTGCTTGCCCTGGGAACTCTGGCACCTTGG 62
```

```
Qy   22 AlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProLysProLysSerAla 41
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63 GCTGTGGAAGGCTCGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCTGC 122

Qy   42 GlnCysLeuArgTyrArgLysProGluCysGlnSerArgTrpGlnCysGlyGlyLysLys 61
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123 CAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTCGGACTGGCAGTGTCCAGGGAAGAAG 182

Qy   62 ArgCysProAspThrCysGlyIleLeuCysLeuAspProValAspThrProAsnPro 81
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  183 AGATGTTGCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAACCCA 242

Qy   82 ThrArgLysProGlyLysCysProValThrTyrGlyIleGlnCysLeuMetLeuAsnPro 101
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  243 ACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGCCAATGTTTGATGCTTAACCCC 302

Qy  102 ProAsnPheCysGluMetAspGlyGlnCysLysArgLeuLysCysGlyMet 121
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  303 CCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGCATG 362

Qy  122 CysGlyLysSerCysValSerProValLysAla 132
        |||||||||||||||||||||||||||||||||
Db  363 TGTGGGAAATCCTGCGTTCCCGTGAAAGCT 395
```

RESULT 27
CD108133
CD108133
LOCUS   CD108133   610 bp   mRNA   linear   EST 15-MAY-2003
DEFINITION AGENCOURT 13979498 NIH_MGC_179 Homo sapiens cDNA clone
            IMAGE:30368824 5', mRNA sequence.
ACCESSION   CD108133
VERSION   CD108133.1 GI:30761307
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 610)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
            cDNA Library Preparation: Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM432 row: d column: 17
            High quality sequence stop: 610.
FEATURES   Location/Qualifiers
     source   1..610
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30368824"
                     /tissue_type="Pituitary"
                     /lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
                     /clone_lib="NIH MGC 179"
                     /note="Organ: brain; Vector: pCMV-SPORT6.1; Site_1: EcoRV
                     (destroyed); Site_2: NotI; Library is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed upon
                     cloning). Average insert size 1.1 kb. Library was
                     constructed by (Invitrogen). Note: this is a NIH_MGC
                     Library."
ORIGIN

Alignment Scores:
Pred. No.:   1.12e-59   Length: 610
Score: 757.00   Matches: 131
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0

Query Match: 99.34%   Indels: 0
DB: 6   Gaps: 0

US-10-613-105-2 (1-132) x CD108133 (1-610)

Qy   2 LysSerGlyLeuPheProPheLeuValLeuLeuLeuGlyThrLeuAlaProTrp 21
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 AAGTCCAGCGGCCTCTTCCCTTCCTTGGTGCTTGCCCTGGGAACTCTGGCACCTTGG 60

Qy   22 AlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysSerAla 41
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GCTGTGGAAGGCTCTGGAAGTCCTTCAAAGCTGGAGTCTGTCCCTCCTAAGAAATCTGCC 120

Qy   42 GlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpProGlnProGlyLysLys 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGTCCAGGGAGAAG 180

Qy   62 ArgCysCysAspThrCysGlyIleLysCysLeuAspProValAspProThrProAsnPro 81
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AGATGTTGTCCTGACACTTGTGGCATCAAATGCCTTGGATCCTGTTGACACCCCAAACCCA 240

Qy   82 ThrArgArgLysProGlyLysLysProValThrTyrGlyGlnCysLeuMetLeuAsnPro 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAACCCC 300

Qy   102 ProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysMetGlyMet 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCCAATTTCTGTGAGATGGATGGGCCAGCGTGCAAGCGTGACTTGAAGTGTTCATGGGCATG 360

Qy   122 CysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||
Db   361 TGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 393

RESULT 28
CB152458
LOCUS   CB152458   612 bp   mRNA   linear   EST 29-JAN-2003
DEFINITION K-EST0209654 C1SNU17 Homo sapiens cDNA clone C1SNU17-39-E09 5',
            mRNA sequence.
ACCESSION   CB152458
VERSION   CB152458.1 GI:28137412
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 612)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE   21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT   Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 39 row: E column: 09
            High quality sequence stop: 612.
FEATURES   Location/Qualifiers
     source   1..612
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="C1SNU17-39-E09"
                     /sex="F"
                     /tissue_type="Uterine"
                     /cell_type="Epithelial"
                     /cell_line="SNU-17"
                     /lab_host="Top10F'"
                     /clone_lib="C1SNU17"
                     /note="Organ: Cervix; Vector: pCNS-D2; Site_1: EcoRI;

Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including
EcoRI site by treatment of T4 RNA ligase and the first
strand cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Alignment Scores:
Pred. No.:   1.13e-59   Length: 612
Score: 757.00   Matches: 131
Percent Similarity: 99.24%   Conservative: 0
Best Local Similarity: 99.24%   Mismatches: 0
Query Match: 99.34%   Indels: 1
DB: Gaps: 0

US-10-613-105-2 (1-132) x CB152458 (1-612)

Qy   1 MetLysSerGlyLeuPheProPheLeuValLeuLeuLeuGlyThrLeuAlaPro 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   23 ATGAAGTCCAGCGGCCTCTTCCCTTCCTTGGTGCTTGCCCTGGGAACTCTGGCACCT 82

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysSer 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   83 TGGGCTGTGGAAGGCTCTGGAAGTCCTTCAAAGCTGGAGTCTGTCCCTCCTAAGAAATCT 142

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpProGlnProGlyLys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   143 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCCAGTGTCCAGGGAAG 202

Qy   61 LysArgCysCysAspThrCysGlyIleLysCysLeuAspProValAspProThrProAsn 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   203 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 262

Qy   81 ProThrArgArgLysProGlyLysLysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 322

Qy   101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysMetGly 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   323 CCCCCCAATTTCTGTGAGATGGATGGGCCAGCGTGCAAGCGTGACTTGAAGTGTTCATGGGC 382

Qy   121 MetCysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||
Db   383 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 418

RESULT 29
BQ691723
LOCUS   BQ691723
DEFINITION AGENCOURT_8046219 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6209402
            5', mRNA sequence.
ACCESSION   BQ691723
VERSION   BQ691723.1 GI:21817039
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 906)
  AUTHORS   Strausberg,R.
  TITLE   NIH-MGC http://mgc.nci.nih.gov/.
  JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov

SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 541)
AUTHORS   Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
            Zeng,Y.-X.
TITLE   Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL   Unpublished (2003)
COMMENT   Contact: YiXin Zeng
            Cancer Center
            Sun Yat-sen University
            651 DongFeng Road East, GuangZhou 510060, China
            Tel: 86-1380-9770-743
            Fax: 86-20-8775-4506
            Email: yxzeng@gzsums.edu.cn.
FEATURES   Location/Qualifiers
     source   1..541
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /tissue_type="normal nasopharynx"
                 /clone_lib="human nasopharynx"
                 /note="ESTs generated from a normal nasopharynx cDNA
                 library from southern Chinese"
ORIGIN

Alignment Scores:
Pred. No.:   1.19e-59   Length: 541
Score: 756.00   Matches: 131
Percent Similarity: 99.24%   Conservative: 0
Best Local Similarity: 99.24%   Mismatches: 1
Query Match: 99.21%   Indels: 0
DB: Gaps: 0

US-10-613-105-2 (1-541) x CD690425 (1-541)

```
Qy     1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
Db    63 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCT 122

Qy    21 TrpAlaValGluGlySerGlyLysPheLysAlaGlyValCysProProLysLysSer  40
Db   123 TGGGCTGTGGAGGGCTCTGGAAAGTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 182

Qy    41 AlaGlnCysLeuArgThrTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
Db   183 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 242

Qy    61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
Db   243 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 302

Qy    81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyIleCysLeuMetLeuAsn 100
Db   303 CCAACAAGGAGGAAGCCTGGAAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 362

Qy   101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
Db   363 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 422

Qy   121 MetCysGlyLysSerCysValSerProValLysAla 132
Db   423 ATGTGTGGAAAATCCTGCGTTTCCCCTGTGACAGCT 458
```

RESULT 31
AA572950
LOCUS   AA572950   582 bp   mRNA   linear   EST 12-SEP-1997
DEFINITION nm33a03.s1 NCI_CGAP_Lip2 Homo sapiens cDNA clone IMAGE:1061932
            similar to gb:X04470 ANTILEUKOPROTEINASE 1 (HUMAN);, mRNA sequence.
ACCESSION   AA572950
VERSION   AA572950.1 GI:2347478
KEYWORDS   EST.

SOURCE
ORGANISM
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2368   row: k   column: 03
            High quality sequence stop: 605.
FEATURES   Location/Qualifiers
     source   1..906
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:6209402"
                 /tissue_type="ductal carcinoma, cell line"
                 /lab_host="DH10B (phage-resistant)"
                 /clone_lib="NIH_MGC_110"
                 /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                 Site_2: EcoRI; cDNA made by oligo-dT priming.
                 Directionally cloned into EcoRI/XhoI sites using the
                 following 5' adaptor: GGCACGAG(G). Library constructed by
                 Ling Hong in the laboratory of Gerald M. Rubin (University
                 of California, Berkeley) using ZAP-cDNA synthesis kit
                 (Stratagene) and Superscript II RT (Life Technologies).
                 Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:   1.89e-59   Length: 906
Score: 757.00   Matches: 131
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 99.34%   Indels: 0
DB: Gaps: 5

US-10-613-105-2 (1-132) x BQ691723 (1-906)

```
Qy     2 LysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaProTrp  21
Db     3 AGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCTTGG  62

Qy    22 AlaValGluGlySerGlyLysPheLysAlaGlyValCysProProLysLysSerAla  41
Db    63 GCTGTGGAGGGCTCTGGAAAGTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCTGCC 122

Qy    42 GlnCysLeuArgThrTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLys  61
Db   123 CAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAGAAG 182

Qy    62 ArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsnPro  81
Db   183 AGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAACCCA 242

Qy    82 ThrArgArgLysProGlyLysCysProValThrTyrGlyIleCysLeuMetLeuAsnPro 101
Db   243 ACAAGGAGGAAGCCTGGAAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAACCCC 302

Qy   102 ProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMet 121
Db   303 CCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGCATG 362

Qy   122 CysGlyLysSerCysValSerProValLysAla 132
Db   363 TGTGGAAAATCCTGCGTTTCCCCTGTGACAGCT 395
```

RESULT 30
CD690425
LOCUS   CD690425   541 bp   mRNA   linear   EST 25-JUN-2003
DEFINITION EST6948 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD690425
VERSION   CD690425.1 GI:32211162
KEYWORDS   EST.

```
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgap8-r@mail.nih.gov
            Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 448   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 437.
FEATURES             Location/Qualifiers
     source          1..582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1061932"
                     /tissue_type="liposarcoma"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Lip2"
                     /note="Vector: pAMP10; mRNA made from liposarcoma, cDNA
                     made by oligo-dT priming; Non-directionally cloned.
                     Size-selected on agarose gel, average insert size 600 bp.
                     Reference: Krizman et al. (1996) Cancer Research
                     56:5380-5383."

ORIGIN
Alignment Scores:
Pred. No.:                  1.31e-59          Length:            582
Score:                      756.00            Matches:           131
Percent Similarity:         99.24%            Conservative:      0
Best Local Similarity:      99.24%            Mismatches:        1
Query Match:                99.21%            Indels:            0
DB:                                           Gaps:              0

US-10-613-105-2 (1-132) x AA572950 (1-582)

Qy      1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaAlaLeuGlyThrLeuAlaPro 20
          |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db      9 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTGGCAGCCCTGGGAACTCTGGCACCT 68

Qy     21 TrpAlaValGluGlyLysGlyLysSerPheLysProProLysLysLysSer 40
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     69 TGGGCTGTGGAAGGCGTCGGAAAGTCCTTCAAAGCTGGAGTCGTGTCCTCCTAAGAAATCT 128

Qy     41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    129 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 188

Qy     61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    189 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 248

Qy     81 ProThrArgArgLysProGlyLysProValThrTyrGlyLeuMetLeuAsn 100
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    249 CCAACAAGGAGGAAGCCTGGGAAGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 308

Qy    101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    309 CCCCCCAATTTCTGTGAGATGGATGGGCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 368

Qy    121 MetCysGlyLysSerCysValSerProValLysAla 132
```

```
Db    369 ATGTGTGGAAATCCTGCGTTCCCGTGAAAGCT 404
          |||||||||||||||||||||||||||||||||

RESULT 32
CD690400
LOCUS       CD690400                  582 bp    mRNA     linear     EST 25-JUN-2003
DEFINITION  EST6923 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD690400
VERSION     CD690400.1  GI:32211112
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
AUTHORS     Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
            Zeng,Y.-X.
TITLE       Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL     Unpublished (2003)
COMMENT     Contact: YiXin Zeng
            Cancer Center
            Sun Yat-sen University
            651 DongFeng Road East, GuangZhou 510060, China
            Tel: 86-1380-9770-743
            Fax: 86-20-8775-4506
            Email: yxzeng@gzsums.edu.cn.
FEATURES             Location/Qualifiers
     source          1..582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="normal nasopharynx"
                     /clone_lib="human nasopharynx"
                     /note="ESTs generated from a normal nasopharynx cDNA
                     library from southern Chinese"

ORIGIN
Alignment Scores:
Pred. No.:                  2e-59             Length:            582
Score:                      754.00            Matches:           131
Percent Similarity:         99.24%            Conservative:      0
Best Local Similarity:      99.24%            Mismatches:        1
Query Match:                98.95%            Indels:            0
DB:                                           Gaps:              0

US-10-613-105-2 (1-132) x CD690400 (1-582)

Qy      1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaAlaLeuGlyThrLeuAlaPro 20
          |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db     62 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTGGCAGCCCTGGGAACTCTGGCACCT 121

Qy     21 TrpAlaValGluGlyLysGlyLysSerPheLysProProLysLysLysSer 40
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 TGGGCTGTGGAAGGCGTCGGAAAGTCCTTCAAAGCTGGAGTCGTGTCCTCCTAAGAAATCT 181

Qy     41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCACGGGAAG 241

Qy     61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 301

Qy     81 ProThrArgArgLysProGlyLysProValThrTyrGlyLeuMetLeuAsn 100
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302 CCAACAAGGAGGAAGCCTGGGAAGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 361

Qy    101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    362 CCCCCCAATTTCTGTGAGATGGATGGGCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 421

Qy    121 MetCysGlyLysSerCysValSerProValLysAla 132
          |||||||||||||||||||||||||||||||||||
```

Db   422 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 457

RESULT 33
BU927919
LOCUS   BU927919   633 bp   mRNA   linear   EST 18-OCT-2002
DEFINITION AGENCOURT_10434154 NIH_MGC_126 Homo sapiens cDNA clone
            IMAGE:6653451 5', mRNA sequence.
ACCESSION   BU927919
VERSION   BU927919.1 GI:24116649
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 633)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: NCI
            cDNA Library Preparation: Michael Brownstein Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2903 row: i column: 03
            High quality sequence stop: 553.
FEATURES   Location/Qualifiers
source   1..633
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:6653451"
             /tissue_type="mixed (pool of 40 RNAs)"
             /lab_host="DH10B (T1-phage-resistant)"
             /clone_lib="NIH_MGC_126"
             /note="Vector: pDNR-LIB; Site 1: SfiI (ggccattggcc);
             Site 2: SfiI (ggccgcctcggcc)_Double-stranded cDNA was
             prepared from a pool of 40 cell line polyA+ RNAs (bladder
             - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
             4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
             kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell -
             5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%,
             salivary gland - 1.3%, and skin 2.3%), 5' and 3'
             adaptors were used in cloning as follows:
             5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGG-3' and
             5'-ATTCTAGAGGCCGAGCGGCCGACATG-dT(30)NN-3'. Full-length
             enriched library was constructed using the Clontech
             Creator SMART kit and size-selected to contain the 0.5-1
             kb size fraction (other fractions present in NIH_MGC_127
             and NIH_MGC_128). Library created in the laboratory of T.
             Usdin, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH_MGC
             Library."

ORIGIN

Alignment Scores:			
Pred. No.:	2.23e-59	Length:	633
Score:	754.00	Matches:	131
Percent Similarity:	99.24%	Conservative:	0
Best Local Similarity:	99.24%	Mismatches:	1
Query Match:	98.95%	Indels:	0
DB:	5	Gaps:	0

US-10-613-105-2 (1-132) x BU927919 (1-633)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuValAlaLeuGlyThrLeuAlaPro 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   25 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGGTCGCCCTGGGAACTCTGGCACCT 84

Qy   21 TrpAlaValGluGluLysGlyLysSerPheLysAlaGlyValCysProProLysLysSer 40
          |||||||||||||||||||||||||||||||||||||||||||||

Db   85 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAAAGTCTGTCCCCTAAGAAATCT 144

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   145 GCCCAGTGCCTTATATACAAGAAACCTGAGTGCCAGTGACTGGCAGTGTCCAGGGAAG 204

Qy   61 LysArgCysProAspThrCysGlyIleLeuCysGlyLeuAspProValAspThrProAsn 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   205 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 264

Qy   81 ProThrThrArgArgLysProGlyLysProValThrThrTyrGlyGlnCysLeuMetLeuAsn 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   265 CCAACAGGAGGAGCCTGGAAGTGCCCAGTGACTTATGGCCAATGTTGATGCTTAAC 324

Qy 101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysMetGly 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   325 CCCCCCAATTCTGTGAGATGGATGGCCAGTGCAAGCCTGACTTGAAGTGTTGCATGGGC 384

Qy 121 MetCysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||||||||||||
Db   385 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 420

RESULT 34
BM694343
LOCUS   BM694343   493 bp   mRNA   linear   EST 28-FEB-2002
DEFINITION UI-E-CI1-afp-h-21-0-UI.r1 UI-E-CI1 Homo sapiens cDNA clone
            UI-E-CI1-afp-h-21-0-UI 5', mRNA sequence.
ACCESSION   BM694343
VERSION   BM694343.1 GI:19007601
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 493)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE   Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED   8889548
COMMENT   Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
            cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
FEATURES   Location/Qualifiers
source   1..493
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="UI-E-CI1-afp-h-21-0-UI"
             /tissue_type="RPE and Choroid"
             /dev_stage="adult"
             /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
             /clone_lib="UI-E-CI1"
             /note="Organ: eye; Vector: pTT3-Pac (Pharmacia) with a
             modified polylinker; Site 1: EcoR I; Site 2: Not I;
             UI-E-CI1 is a normalized cDNA library containing the
             following tissue(s): RPE and Choroid. The library was
             constructed according to Bonaldo, Lennon and Soares,
             Genome Research, 6:791-806, 1996. First strand cDNA
             synthesis was primed with an oligo-dT primer containing a
             Not I site. Double stranded cDNA was ligated to an EcoR I

```
                    adaptor, digested with Not I, and cloned directionally
                    into pT7T3-Pac vector. The oligonucleotide used to prime
                    the synthesis of first-strand cDNA contains a library tag
                    sequence that is located between the Not I site and the
                    (dT)18 tail. The sequence tag for this library is ACCTA.
                    This library was created for the program, Gene Discovery
                    in the Visual System, supported by National Eye Institute
                    (NEI)."

ORIGIN

Alignment Scores:
Pred. No.:            5.73e-59          Length:         493
Score:                748.00            Matches:        129
Percent Similarity:   100.00%           Conservative:   0
Best Local Similarity: 100.00%          Mismatches:     0
Query Match:          98.16%            Indels:         0
DB:                   4                 Gaps:           0

US-10-613-105-2 (1-132) x BM694343 (1-493)

Qy    4 SerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaProTrpAlaVal 23
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3 AGCGGGCCTCTTCCCCTTCCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCTTGGGCTGTG 62

Qy   24 GluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSerAlaGlnCys 43
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63 GAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCTGCCCAGTGC 122

Qy   44 LeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCys 63
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123 CTTAGATACAAGAAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAGAAGAGATGT 182

Qy   64 CysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArg 83
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  183 TGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAACCCAACAAGG 242

Qy   84 ArgLysProGlyLysProValThrTyrGlyGlnCysLeuMetLeuAsnProProAsn 103
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  243 AGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAACCCCCCAAT 302

Qy  104 PheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMetCysGly 123
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  303 TTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGCATGTGTGGG 362

Qy  124 LysSerCysValSerProValLysAla 132
         |||||||||||||||||||||||||||
Db  363 AAATCCTGCGTTTCCCCTGTGAAAGCT 389


RESULT 35
BM704851
LOCUS       BM704851                 567 bp    mRNA        linear    EST 28-FEB-2002
DEFINITION  UI-E-CI1-afp-h-21-0-UI.r2 UI-E-CI1 Homo sapiens cDNA clone
            UI-E-CI1-afp-h-21-0-UI 5', mRNA sequence.
ACCESSION   BM704851
VERSION     BM704851.1  GI:19018109
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 567)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   MEDLINE   97044477
   PUBMED    8889548
  COMMENT   Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road, 4156    MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
```

```
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.

FEATURES             Location/Qualifiers
     source          1..567
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-CI1-afp-h-21-0-UI"
                     /tissue_type="RPE and Choroid"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-CI1"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site 1: EcoR I; Site 2: Not I;
                     UI-E-CI1 is a normalized cDNA library containing the
                     following tissue(s): RPE and Choroid. The library was
                     constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806, 1996. First strand cDNA
                     synthesis was primed with an oligo-dT primer containing a
                     Not I site. Double stranded cDNA was ligated to an EcoR I
                     adaptor, digested with Not I, and cloned directionally
                     into pT7T3-Pac vector. The oligonucleotide used to prime
                     the synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is ACCTA.
                     This library was created for the program, Gene Discovery
                     in the Visual System, supported by National Eye Institute
                     (NEI)."

ORIGIN

Alignment Scores:
Pred. No.:            6.88e-59          Length:         567
Score:                748.00            Matches:        129
Percent Similarity:   100.00%           Conservative:   0
Best Local Similarity: 100.00%          Mismatches:     0
Query Match:          98.16%            Indels:         0
DB:                                     Gaps:           0

US-10-613-105-2 (1-132) x BM704851 (1-567)

Qy    4 SerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaProTrpAlaVal 23
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3 AGCGGGCCTCTTCCCCTTCCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCTTGGGCTGTG 62

Qy   24 GluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSerAlaGlnCys 43
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63 GAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCTGCCCAGTGC 122

Qy   44 LeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCys 63
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123 CTTAGATACAAGAAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAGAAGAGATGT 182

Qy   64 CysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArg 83
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  183 TGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAACCCAACAAGG 242

Qy   84 ArgLysProGlyLysProValThrTyrGlyGlnCysLeuMetLeuAsnProProAsn 103
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  243 AGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAACCCCCCAAT 302

Qy  104 PheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMetCysGly 123
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  303 TTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGCATGTGTGGG 362

Qy  124 LysSerCysValSerProValLysAla 132
         |||||||||||||||||||||||||||
Db  363 AAATCCTGCGTTTCCCCTGTGAAAGCT 389
```

```
RESULT 36
BI489624
LOCUS       BI489624                962 bp    mRNA    linear   EST 28-AUG-2001
DEFINITION  603032084F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173293 5',
            mRNA sequence.
ACCESSION   BI489624
VERSION     BI489624.1  GI:15328852
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 962)
AUTHORS     NIH-MGC http://mgc.ncl.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11431  row: g  column: 22
            High quality sequence stop: 555.
FEATURES             Location/Qualifiers
     source          1..962
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5173293"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC 115"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
                     source anonymous pool of 6 male brains, age range 23-27; 1
                     male lung, age 27; and 1 male testis, age 69. Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon cloning). Average insert size 1.8 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was  constructed by C.
                     Gruber (Invitrogen).  Research Genetics tracking code
                     021. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.:             1.38e-58         Length:         962
Score:                 748.00           Matches:        129
Percent Similarity:    100.00%          Conservative:   0
Best Local Similarity: 100.00%          Mismatches:     0
Query Match:           98.16%           Indels:         0
DB:                                     Gaps:           0

US-10-613-105-2 (1-132) x BI489624 (1-962)

Qy     4 SerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaProTrpAlaVal 23
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    18 AGCGGCCTCTTCCCCTTCCTGGTGCTGCTGGCCCTCGGAACTCTGGCACCTTGGGCTGTG 77

Qy    24 GluGlyGlySerGlyLysPheLysSerPheLysAlaGlyValValCysProLysSerAlaGlyGlnCys 43
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    78 GAAGGCTCTCGAAAGTCCTTCAAAGCTGGAGTCGTGTCCTCCTAAGAAATCTGCCCAGTGC 137

Qy    44 LeuArgTyrLysLysProGluCysGlnGlnSerAspTrpGlnCysProGlyLysLysArgCys 63
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   138 CTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAGAAGAGATGT 197

Qy    64 CysProAspThrCysGlyIleLysCysProGlyLeuLysLeuAspProValAspProThrArg 83
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   198 TGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAGG 257
```

```
Qy    84 ArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsnProProAsn 103
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   258 AGGAAGCCTGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAACCCCCCAAT 317

Qy   104 PheCysGluMetAspGlyGlyGlnCysLysLysArgAspLeuLysCysSerMetGlyMetCysGly 123
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   318 TTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGCATGTGTGG 377

Qy   124 LysSerCysValSerProValLysAla 132
           ||||||||||||||||||||||||
Db   378 AAATCCTGCGTTTCCCCTGTGAAAGCT 404


RESULT 37
AW264225/c
LOCUS       AW264225                575 bp    mRNA    linear   EST 28-DEC-1999
DEFINITION  xq99c06.x1 NCI_CGAP_Brn53 Homo sapiens cDNA clone IMAGE:2758762 3'
            similar to gb:X04470 ANTILEUKOPROTEINASE 1 (HUMAN);, mRNA sequence.
ACCESSION   AW264225
VERSION     AW264225.1  GI:6641126
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 575)
AUTHORS     NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
JOURNAL     Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life
            Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
            I.M.A.G.E. Consortium DNA Sequencing by: Washington University
            Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www.bio.llnl.gov/bbrp/image/iMage.html

            Possible reversed clone: polyT not found
            Seq primer: -40UP from Gibco
            High quality sequence stop: 411.
FEATURES             Location/Qualifiers
     source          1..575
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2758762"
                     /tissue_type="three pooled meningiomas"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Brn53"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Library constructed by Life Technologies."

ORIGIN

Alignment Scores:
Pred. No.:             8.67e-59         Length:          575
Score:                 747.00           Matches:         129
Percent Similarity:    98.47%           Conservative:    0
Best Local Similarity: 98.47%           Mismatches:      2
Query Match:           98.03%           Indels:          0
DB:                                     Gaps:            0

US-10-613-105-2 (1-132) x AW264225 (1-575)

Qy     2 LysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaProTrp 21
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   573 AGTTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTGGCCCTCGGGGAACTCTGGCACCTTGG 514
```

EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

```
Alignment Scores:
Pred. No.:
Score:                  1.08e-58      Length:        580
Percent Similarity:     98.48%        Matches:       746.00
Best Local Similarity:  98.48%        Conservative:  0
Query Match:            97.90%        Mismatches:    2
DB:                                   Indels:        0
                                      Gaps:          0
```

US-10-613-105-2 (1-132) x CB151317 (1-580)

```
Qy    1 MetLysSerGlyLysLeuProPheProLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   37 ATGAAGAGTGGAAAGCTTCCTTTCCCCTTGTGCTGCTGGCTGGAACCTGGCACCT      96

Qy   21 TrpAlaValGlySerGlyLysSerPheLysValCysProProLysSer              40
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   97 TGGGCTGTGGGAAGGCTCTGGGAAAGTCCTTCAAAGCTGGAGTCGTCCTAATAAATCT   156

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysCysGlnSerAspTrpGlnCysProGlyLys  60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  157 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTGACTGGCAGTGTCCAGGGAAG    216

Qy   61 LysArgCysCysProAspThrCysGlyIleIleLysCysLeuAspProValAspThrProAsn  80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  217 AAGAGATGTTACCCGACCACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC  276

Qy   81 ProThrArgArgLysProGlyLysCysProValThrThrTyrGlyGlnCysLeuMetLeuAsn 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  277 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTGATGCTTAAC   336

Qy  101 ProProAsnPheCysGluMetAspAspGlyGlnCysLysGlyArgAspLeuLysCysMetGly 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  337 CCCCCAATTTCTGTGAGATGGATGGCCAAGCGTGACTTGAAGTGTGCATGGGC        396

Qy  121 MetCysGlyLysSerCysValSerProValLysAla                          132
          ||||||||||||||||||||||||||||||||||||
Db  397 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT                          432
```

```
Qy   22 AlaValGluGlySerGlyLysPheLysAlaGlyValCysProProLysSerAla         41
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  513 GCTGTGGGAAGGCTCTGGGAAAGTCCTTCAAAGCTGGAGTCGTCCTAAGAAATCTGCC   454

Qy   42 GlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLys   61
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  453 CAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAGAAG  394

Qy   62 ArgCysCysProAspThrCysGlyIleIleLysCysLeuAspProValAlaProThrProAsnPro 81
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  393 AGATGTTGTCCTGACACCTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAACCCA  334

Qy   82 ThrArgArgLysProGlyLysCysProValThrThrTyrGlyGlnCysLeuMetLeuAsnPro 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  333 ACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTACTTATGGCCAATGTTGATGCTTAACCCC 274

Qy  102 ProAsnPheCysGluMetAspAspGlyGlnCysLysArgAspAspLeuLysCysMetGlyMet 121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  273 CCCAATTTCTGTGAGATGGATGGCCAAGCGTGACTTGAAGTGTTGCATGGGCATG        214

Qy  122 CysGlyLysSerCysValSerProValLysAla                             132
          |||||||||||||||||||||||||||||||||
Db  213 TGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT                             181
```

RESULT 38
CB151317

```
LOCUS       CB151317       580 bp    mRNA      linear    EST 29-JAN-2003
DEFINITION  K-EST0208192 C1SNU17 Homo sapiens cDNA clone C1SNU17-28-C11 5',
            mRNA sequence.
ACCESSION   CB151317
VERSION     CB151317.1  GI:28135601
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 28 row: C column: 11
            High quality sequence stop: 580.
FEATURES             Location/Qualifiers
     source          1..580
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="C1SNU17-28-C11"
                     /sex="F"
                     /tissue_type="Uterine"
                     /cell_type="Epithelial"
                     /cell_line="SNU-17"
                     /lab_host="Top10F'"
                     /clone_lib="C1SNU17"
                     /note="Organ: Cervix; Vector: pCNS-D2; Site 1: EcoRI;
                     Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including
                     EcoRI site by treatment of T4 RNA ligase and the first
                     strand cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
```

RESULT 39
BE396187

```
LOCUS       BE396187       483 bp    mRNA      linear    EST 21-JUL-2000
DEFINITION  601308817F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3626789 5',
            mRNA sequence.
ACCESSION   BE396187
VERSION     BE396187.1  GI:9341552
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 483)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/,
            National Institutes of Health, Mammalian Gene Collection (MGC)
  TITLE     Unpublished (1999)
  JOURNAL
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM309 row: b column: 06
            High quality sequence start: 5
```

COMMENT   Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM305   row: h   column: 03
High quality sequence start: 4
High quality sequence stop: 560.

FEATURES   Location/Qualifiers
source   1..581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3625394"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_44"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:
Pred. No.:   Length: 581
Score: 5.92e-58   Matches: 127
Percent Similarity: 738.00   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 96.85%   Gaps: 0

US-10-613-105-2 (1-132) x BE392080 (1-581)

Qy   6 LeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaProTrpAlaValGluGly 25
Db   1 CTCTTCCCCTTCCTGGTGCTGCTGGCCCTGGGAACTCTGGCACCTTGGGCTGTGGAAGGC 60

Qy 26 SerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeuArg 45
Db 61 TCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCTGCCCAGTGCCTTAGA 120

Qy 46 TyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysPro 65
Db 121 TACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAGAAGAGATGTTGTCCT 180

Qy 66 AspThrCysGlyIleLysCysLeuAspProValAspAspThrProAsnProThrArgArgLys 85
Db 181 GACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAACCCAACAGGAGGAAG 240

Qy 86 ProGlyLysCysProValThrTyrGlyValGlnCysLeuMetLeuAsnProAsnPheCys 105
Db 241 CCTGGGAAGTGCCCAGTGACTTATGGCCAAATGTTGATGCTTAACCCCCAATTTCTGT 300

Qy 106 GluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMetCysGlyLysSer 125
Db 301 GAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGCATGTGTGGGAAATCC 360

Qy 126 CysValSerProValLysAla 132
Db 361 TGCGTTTCCCCTGTGAAAGCT 381

RESULT 41
BE743613
LOCUS   BE743613   557 bp   mRNA   linear   EST 15-SEP-2000
DEFINITION 601573751F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3834636 5',
mRNA sequence.

FEATURES   High quality sequence stop: 481.
source   Location/Qualifiers
1..483
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3626789"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_44"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:
Pred. No.:   Length: 483
Score: 3.04e-58   Matches: 128
Percent Similarity: 740.00   Conservative: 0
Best Local Similarity: 98.46%   Mismatches: 2
Query Match: 98.46%   Indels: 0
DB: 97.11%   Gaps: 0

US-10-613-105-2 (1-132) x BE396187 (1-483)

Qy   3 SerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaProTrpAla 22
Db 34 AGTCAGCGGCTTCTTCCCCTTCCTGGTGCTGCTGGCCCTGGGAACTCTGGCACCTTGGGCT 93

Qy 23 ValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGln 42
Db 94 GTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCTGCCCAG 153

Qy 43 CysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArg 62
Db 154 TGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAGAAGAGA 213

Qy 63 CysCysProAspThrCysGlyIleLysCysLeuAspProValAspAspThrProThr 82
Db 214 TGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAACCCAACA 273

Qy 83 ArgArgLysProGlyLysCysProValThrTyrGlyValGlnCysLeuMetLeuAsnProPro 102
Db 274 AGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAAATGTTTGATGCTTAACCCCCC 333

Qy 103 AsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMetCys 122
Db 334 AATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGCATGTGT 393

Qy 123 GlyLysSerCysValSerProValLysAla 132
Db 394 GGGAAATCCTGCGTTTCCCCTGTGAAAGCT 423

RESULT 40
BE392080
LOCUS   BE392080   581 bp   mRNA   linear   EST 21-JUL-2000
DEFINITION 601307712F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3625394 5',
mRNA sequence.
ACCESSION   BE392080
VERSION   BE392080.1 GI:9337445
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 581)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)

ACCESSION   BE743613
VERSION   BE743613.1 GI:10157605
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 557)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE   Unpublished (1999)
JOURNAL
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM515 row: f column: 13
            High quality sequence stop: 557.
FEATURES   Location/Qualifiers
     source   1..557
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3834636"
                     /tissue_type="adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_9"
                     /note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACCAG(G). Size-selected >500bp for average
                     insert size 1.8kb. Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Supercript II RT (Life Technologies)."

ORIGIN

Alignment Scores:
Pred. No.:
Score: 1.3e-57   Length: 557
Percent Similarity: 734.00   Matches: 126
Best Local Similarity: 100.00%   Conservative: 0
Query Match: 100.00%   Mismatches: 0
DB: 96.33%   Indels: 0
                                    Gaps: 2

US-10-613-105-2 (1-132) x BE743613 (1-557)

Qy   7 PheProPheLeuValLeuAlaLeuGlyThrLeuAlaProTrpAlaValGluGlySer   26
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2 TTCCCCTTCCTGGTCGTGCTGCTCTGGGACTCTGGCCACCTTGGCGTGTGGAAGGCTCT 61

Qy   27 GlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeuArgTyr   46
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 GGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCTGCCCAGTGCCTTAGATAC 121

Qy   47 LysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysProAsp 66
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122 AAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAGAAGAGATGTTGTCCTGAC 181

Qy   67 ThrCysGlyIleLeuLysCysLeuAspProValAspThrProAsnProThrArgArgLysPro 86
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182 ACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAACCCAACAAGGAGGAAGCCT 241

Qy   87 GlyLysCysProValThrTyrGlyIleGlnCysLeuMetLeuAsnProProAsnPheCysGlu 106
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 242 GGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAACCCCCCAAATTTCTGTGAG 301

Qy 107 MetAspAlaGlyGlnCysProArgAspAspLeuLysCysCysMetGlyMetCysGlyLysSerCys 126
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 302 ATGGATGCGGGCCAAGCGTGCAAGCCGATGACTTGAAGTGTTGCATGGGCATGTGTGGGAAATCCTGC 361

Qy 127 ValSerProValLysAla 132
        ||||||||||||||||||
Db 362 GTTTCCCCTGTGAAAGCT 379

RESULT 42
BI819626
LOCUS   BI819626   583 bp   mRNA   linear   EST 04-OCT-2001
DEFINITION 603034162F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175327 5',
            mRNA sequence.
ACCESSION   BI819626
VERSION   BI819626.1 GI:15931176
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 583)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE   Unpublished (1999)
JOURNAL
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11436 row: 1 column: 16
            High quality sequence stop: 562.
FEATURES   Location/Qualifiers
     source   1..583
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5175327"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_115"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 6 male_brains, age range 23-27; 1
                     male lung, age 27; and 1 male testis, age 69. Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon cloning). Average insert size 1.8 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was constructed by C.
                     Gruber (Invitrogen). Research Genetics tracking code
                     021. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.:
Score: 1.39e-57   Length: 583
Percent Similarity: 734.00   Matches: 126
Best Local Similarity: 100.00%   Conservative: 0
Query Match: 100.00%   Mismatches: 0
DB: 96.33%   Indels: 0
                                    Gaps: 4

US-10-613-105-2 (1-132) x BI819626 (1-583)

Qy   7 PheProPheLeuValLeuAlaLeuGlyThrLeuAlaProTrpAlaValGluGlySer   26
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2 TTCCCCTTCCTGGTGCTGCTGCTCTGGGAACTCTGGCACCTTGGCGTGTGGAAGGCTCT 61

Qy   27 GlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeuArgTyr   46
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 GGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCTGCCCAGTGCCTTAGATAC 121

Qy   47 LysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysProAsp 66
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122 AAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAGAAGAGATGTTGTCCTGAC 181

```
Qy     1  MetLysSerSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro  20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    23  ATGAAGTCCAGCGGCCTCTTCCCCTTCCCCGTGCTGGCTCTGGGAACCCTGGCACCT  82

Qy    21  TrpAlaValGluGlySerGlyValSerPheLysAlaGlyValCysProProLysSer  40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    83  TGGGCTGTGGAAGGCTCTGGAAGTCCTTCAAAGCTGGAGTCTGTCTCCTAAGAAATCT  142

Qy    41  AlaGlnCysLeuArgTyrLArgLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   143  GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTGACTGGCAGTGTCCAGGGAAG  202

Qy    61  LysArgCysProAspThrCysLeuLysLysCysProValThrTyrGlyLeuMetLeuAsn  80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   203  AAGAGATGTTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAA-  261

Qy    81  ProThrArgArgLysProGlyLysCysProValThrTyrGlyLeuMetLeuAsn  100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   262  CCAACAAGGAGGAAGCCTGGAAGTGCCCAGTGACTTATGGCCAATGTTGATGCTTAAC  321

Qy   101  ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   322  CCCCCCAATTTCTGTGAGATGGATGGGCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC  381

Qy   121  MetCysGlyLysSerCysValSerProValLysAla  132
          ||||||||||||||||||||||||||||||||||||
Db   382  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  417
```

```
RESULT 44
BE271401
LOCUS        BE271401                    793 bp     mRNA        linear      EST 26-OCT-2000
DEFINITION   601140427F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3049710 5',
             mRNA sequence.
ACCESSION    BE271401
VERSION      BE271401.1   GI:9145136
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 793)
  AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
  TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL    Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
             Plate: LLCM94 row: e  column: 07
             High quality sequence stop: 577.
FEATURES             Location/Qualifiers
     source          1..793
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3049710"
                     /tissue_type="adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_9"
                     /note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G). Size-selected >500bp for average
                     insert size 1.8kb. Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN
```

```
Qy    67  ThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArgArgLysPro  86
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182  ACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAACCCAACAAGGAGGAAGCCT  241

Qy    87  GlyLysLysProValThrTyrGlyIleCysLeuMetLeuAsnProProAsnPheCysGlu  106
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242  GGGAAGAAGTGCCCAGTGACTTATGGCCAATGTTGATGCTTAACCCCCCCAATTTCTGTGAG  301

Qy   107  MetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMetCysGlyLysSerCys  126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302  ATGGATGGGCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGCATGTGTGGGAAATCCTGC  361

Qy   127  ValSerProValLysAla  132
          ||||||||||||||||||
Db   362  GTTTCCCCTGTGAAAGCT  379
```

```
RESULT 43
BG484664
LOCUS        BG484664                    770 bp     mRNA        linear      EST 21-MAR-2001
DEFINITION   602505842F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4619099 5',
             mRNA sequence.
ACCESSION    BG484664
VERSION      BG484664.1   GI:13416943
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 770)
  AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
  TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL    Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCM1378 row: d  column: 12
             High quality sequence stop: 554.
FEATURES             Location/Qualifiers
     source          1..770
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4619099"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_77"
                     /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgctcggcc); Site_2: SfiI (ggccattatgcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCGGCCGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C, G, or T). Average insert size 1.9
                     kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."
ORIGIN
```

```
Alignment Scores:
Pred. No.:                           Length:            770
Score:           2.47e-57            Matches:           733.00
Percent Similarity:  99.24%          Conservative:      0
Best Local Similarity:  99.24%       Mismatches:        0
Query Match:         96.19%          Indels:            1
DB:                                  Gaps:              0

US-10-613-105-2 (1-132) x BG484664 (1-770)
```

```
Alignment Scores:
Pred. No.:                                Length:          793
Score:                  7.42e-57          Matches:         125
Percent Similarity:     728.00            Conservative:      0
Best Local Similarity:  100.00%           Mismatches:        0
Query Match:            100.00%           Indels:            0
                                          Gaps:              0

US-10-613-105-2 (1-132) x BE271401 (1-793)

Qy    8 ProPheValLeuLeuAlaLeuGlyThrLeuAlaProTrpAlaValAlaGlyGlySerGly 27
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2 CCCTTCCTGTGCTTGCTGGCACTGGGAACTCTGGCCACCTGGGCTGTGGGAGGCTCTGGA 61

Qy   28 LysSerPheLysValGlyAlaGlnCysLeuArgTyrLys 47
          |||||||||||||||||||||||||||||||||||||||
Db   62 AAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCTGCCCAGTGCCTTAGATACAAG 121

Qy   48 LysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysSerProAspThr 67
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122 AAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAGAAGAGATGTTCCTGACACT 181

Qy   68 CysGlyIleIleLysCysLeuAspAspProValAspThrProAsnProThrArgArgLysProGly 87
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  182 TGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAACCCAACAAGGAGGAAGCCTGGG 241

Qy   88 LysCysProValThrTyrGlyGlnCysLeuMetLeuAsnProProAsnPheCysGluMet 107
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  242 AAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAACCCCCCAAATTTCTGTGAGATG 301

Qy  108 AspGlyTyrGlnCysLysLysLeuLysCysCysMetGlyMetCysGlyLysSerCysVal 127
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  302 GATGGCTACCAGTGCAAGCCGTGACTTGGAAGTGTAGCATGGGCATGTGTGGGAAATCCTGCGTT 361

Qy  128 SerProValLysAla 132
          |||||||||||||||
Db  362 TCCCCGTGAAAGCT 376
```

```
Alignment Scores:
Pred. No.:                                Length:          511
Score:                  6.37e-57          Matches:         130
Percent Similarity:     726.00            Conservative:      0
Best Local Similarity:  97.74%            Mismatches:        2
Query Match:            97.74%            Indels:            1
DB:                     95.28%            Gaps:              0

US-10-613-105-2 (1-132) x CD699640 (1-511)

Qy    1 MetLysSerSerGlyLeuPheProPheValLeuProPheLeuValLeuAlaLeuGlyThrLeuAlaLeuAlaPro 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   44 ATGAAGTCCAGCGGCCTCTTCCCCTTCGTGCTGCCTTGCCTGCTGGCCCTGGGAACTCTGGCCACCT 103

Qy   21 Trp-AlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSe 40
          |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  104 TGTGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATC 163

Qy   40 rAlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLy 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  164 TGCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAA 223

Qy   60 sLysArgCysSerProAspThrCysGlyIleIleLysCysLeuAspAspProValAspThrProAs 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  224 GAAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAA 283

Qy   80 nProThrArgArgLysProGlyTyrGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAs 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  284 CCCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAA 343

Qy  100 nProProAsnPheCysGluMetAspGlyTyrGlnCysLysLysLeuLysCysCysMetGl 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  344 CCCCCCAAATTTCTGTGAGATGGATGGCTACCAGTGCAAGCCGTGACTTGGAAGTGTAGCATGGG 403

Qy  120 yMetCysGlyLysSerCysValSerProValLysSerAla 132
          ||||||||||||||||||||||||||||||||||||
Db  404 CATGTGGGAAATCCTGCGTATCCCCGTGAAAGCT 440
```

Search completed: October 24, 2004, 05:52:24
Job time : 2736 secs

```
RESULT 45
CD699640
LOCUS       CD699640                 511 bp    mRNA    linear    EST 25-JUN-2003
DEFINITION  EST16036 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD699640
VERSION     CD699640.1  GI:32228866
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 511)
AUTHORS     Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
            Zeng,Y.-X.
TITLE       Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL     Unpublished (2003)
COMMENT     Contact: Yixin Zeng
            Cancer Center
            Sun Yat-sen University
            651 DongFeng Road East, Guangzhou 510060, China
            Tel: 86-1380-9770-743
            Fax: 86-20-8775-4506
            Email: yxzeng@zsums.edu.cn.
FEATURES             Location/Qualifiers
     source          1..511
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="normal nasopharynx"
                     /clone_lib="human nasopharynx"
                     /note="ESTs generated from a normal nasopharynx cDNA
                     library from southern Chinese"
ORIGIN
```

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 24, 2004, 04:00:14 ; Search time 427 Seconds
(without alignments)
1583.976 Million cell updates/sec

Title: US-10-613-105-2
Perfect score: 762
Sequence: 1 MKSSGLFPFLVILLALGTLAP........RDLKCCMGMCGKSCVSPVKA 132

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext   0.5
Ygapop 10.0 , Ygapext   0.5
Fgapop   6.0 , Fgapext   7.0
Delop   6.0 , Delext   7.0

Searched: 3407233 seqs, 256960514 residues

Total number of hits satisfying chosen parameters: 6814466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US10613105/runat_22102004_170442_18724/app_query.fasta_1.327
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=100 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=45 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10613105 @CGN_1_1_480_@runat_22102004_170442_18724
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   Published Applications NA:*
1: /cgn2_6/ptcodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptcodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptcodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptcodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptcodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptcodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptcodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptcodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptcodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptcodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptcodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptcodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptcodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptcodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptcodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptcodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptcodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptcodata/2/pubpna/US10_PUBCOMB.seq:*
19: /cgn2_6/ptcodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptcodata/2/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptcodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	762	100.0	399	15	US-10-257-021-53	Sequence 53, App1
2	762	100.0	399	15	US-10-172-118-864	Sequence 864, App
3	762	100.0	399	16	US-10-342-887-864	Sequence 864, App
4	762	100.0	422	16	US-10-309-290-141	Sequence 141, App
5	762	100.0	565	15	US-10-360-368-1	Sequence 1, App1i
6	762	100.0	594	9	US-09-964-824A-582	Sequence 582, App
7	762	100.0	594	9	US-09-954-456-1989	Sequence 1989, Ap
8	762	100.0	594	9	US-09-865-812-1	Sequence 1, App1i
9	762	100.0	594	14	US-10-097-340-295	Sequence 295, App
10	762	100.0	594	15	US-10-205-823-374	Sequence 374, App
11	762	100.0	594	15	US-10-177-293-429	Sequence 429, App
12	762	100.0	594	15	US-10-172-118-863	Sequence 863, App
13	762	100.0	594	16	US-10-305-720-1271	Sequence 1271, Ap
14	762	100.0	594	16	US-10-309-290-143	Sequence 143, App
15	762	100.0	594	16	US-10-240-425-1529	Sequence 1529, App
16	762	100.0	594	16	US-10-342-887-863	Sequence 863, App
17	762	100.0	594	17	US-10-613-105-1	Sequence 1, App1i
18	762	100.0	594	17	US-10-322-696-125	Sequence 125, App
19	762	100.0	598	15	US-10-177-293-431	Sequence 431, App
20	762	100.0	598	16	US-10-173-999-153	Sequence 153, App
21	762	100.0	1084	10	US-09-814-353-20587	Sequence 20587, A
22	638.5	83.8	1525	14	US-10-025-514-15	Sequence 15, App1
23	635	83.3	321	16	US-10-309-290-143	Sequence 3, App1i
24	635	83.3	321	16	US-10-309-290-145	Sequence 145, App
25	635	83.3	1525	14	US-10-025-514-7	Sequence 7, App1i
26	537.5	70.5	22324	17	US-10-322-696-124	Sequence 124, App
27	463.5	60.8	1079	17	US-10-322-696-122	Sequence 122, App
28	458	60.1	325	16	US-10-305-720-1024	Sequence 1024, Ap
29	352.5	46.3	422	15	US-10-125-159-37	Sequence 37, App
30	339	44.5	805	14	US-10-116-802-431	Sequence 431, App
31	285	37.4	4090	14	US-10-076-816-57	Sequence 57, App1
32	283.5	37.2	313	10	US-09-814-353-2826	Sequence 2826, Ap
33	283.5	37.2	313	10	US-09-814-353-9156	Sequence 9156, App
34	277.5	36.4	38692	17	US-10-322-696-121	Sequence 121, App
c 35	264	34.6	411	9	US-09-964-824A-495	Sequence 495, App
c 36	264	34.6	411	9	US-09-954-456-1987	Sequence 1987, Ap
c 37	264	34.6	411	9	US-09-967-768A-3	Sequence 3, App1i
c 38	264	34.6	411	10	US-09-960-706-16	Sequence 16, App1
39	264	34.6	411	10	US-09-870-319-9	Sequence 9, App1i
40	257.5	33.8	248	17	US-10-430-201-2236	Sequence 2236, Ap
41	257.5	33.8	248	17	US-10-430-201-2237	Sequence 2237, Ap
42	243	31.9	188	10	US-09-814-353-15540	Sequence 15540, A
43	232.5	30.5	518	15	US-10-125-159-43	Sequence 43, App1
44	230	30.2	208	15	US-10-125-159-41	Sequence 41, App1
45	224	29.4	292	9	US-09-884-441-377	Sequence 377, App
46	224	29.4	292	10	US-09-907-969-377	Sequence 377, App
47	224	29.4	292	15	US-09-827-271-377	Sequence 377, App
48	224	29.4	292	15	US-09-198-053-377	Sequence 377, App
49	212	27.8	732	15	US-10-221-097-14	Sequence 14, App1
50	207	27.2	558	15	US-10-221-097-15	Sequence 15, App1
51	185.5	24.3	1010	17	US-10-729-807-27	Sequence 27, App1
52	185.5	24.3	1055	16	US-10-250-959-1	Sequence 1, App1i
53	165.5	21.7	374	15	US-10-233-150-12	Sequence 12, App1
54	165.5	21.7	390	15	US-10-233-150-3	Sequence 3, App1i
55	165.5	21.7	486	15	US-10-233-150-10	Sequence 10, App1
56	165.5	21.7	622	9	US-09-925-300-421	Sequence 421, App
57	165.5	21.7	622	16	US-10-264-049-627	Sequence 627, App
58	165.5	21.7	643	10	US-09-992-600A-41	Sequence 41, App1i
59	165.5	21.7	643	10	US-09-924-340-41	Sequence 41, App1i
60	165.5	21.7	643	10	US-09-992-095B-41	Sequence 41, App1i
61	165.5	21.7	643	10	US-09-999-570-41	Sequence 41, App1i
62	165.5	21.7	643	14	US-10-000-489-41	Sequence 41, App1i
63	165.5	21.7	643	14	US-10-000-986-41	Sequence 41, App1i
64	165.5	21.7	643	15	US-10-154-678-41	Sequence 41, App1i
65	165.5	21.7	643	15	US-10-001-142-41	Sequence 41, App1i
66	165.5	21.7	656	14	US-10-198-846-11505	Sequence 11505, A
67	165.5	21.7	1004	16	US-10-264-049-628	Sequence 628, App
68	165.5	21.7	1077	15	US-10-233-150-4	Sequence 4, App1i
69	165.5	21.7	1098	15	US-10-233-150-6	Sequence 6, App1i
70	163	21.4	584	15	US-10-295-027-501	Sequence 501, App

71	163	21.4	584	US-10-173-999-61	Sequence 61, Appl
72	161	21.1	239	US-10-242-535A-1975	Sequence 1975, Ap
73	161	21.1	239	US-10-085-783A-1975	Sequence 1975, Ap
74	161	21.1	583	US-09-967-768A-214	Sequence 214, App
75	161	21.1	583	US-10-097-340-360	Sequence 360, App
76	161	21.1	583	US-10-233-150-8	Sequence 8, Appli
77	161	21.1	583	US-10-177-293-499	Sequence 499, App
78	161	21.1	583	US-10-172-118-1242	Sequence 1242, Ap
79	161	21.1	583	US-10-342-887-1242	Sequence 1242, Ap
80	160.5	21.1	652	US-09-971-429B-55	Sequence 55, Appl
81	158	20.7	5690	US-09-764-891-9939	Sequence 9939, Ap
c 82	154	20.2	351	US-09-867-701-8288	Sequence 8288, Ap
83	152.5	20.0	478	US-09-641-643-1233	Sequence 1233, Ap
84	152.5	20.0	571	US-09-969-034-2513	Sequence 2513, Ap
85	152.5	20.0	582	US-10-157-031-119	Sequence 119, App
86	152.5	20.0	631	US-10-295-027-335	Sequence 335, App
87	149.5	19.6	478	US-10-325-745-11	Sequence 11, Appl
c 88	149	19.6	480	US-09-920-300A-1393	Sequence 1393, Ap
c 89	149	19.6	480	US-10-033-528-1393	Sequence 1393, Ap
c 90	149	19.6	480	US-10-099-926-1393	Sequence 1393, Ap
91	149	19.6	2309	US-09-954-456-1126	Sequence 1126, Ap
92	149	19.6	2309	US-09-954-456-1798	Sequence 1798, Ap
93	149	19.6	2309	US-10-305-720-1249	Sequence 1249, Ap
94	149	19.6	2309	US-10-240-425-1277	Sequence 1277, Ap
95	146.5	19.2	762	US-09-989-722-344	Sequence 344, App
96	146.5	19.2	762	US-09-989-723-344	Sequence 344, App
97	146.5	19.2	762	US-09-989-279-344	Sequence 344, App
98	146.5	19.2	762	US-09-989-727-344	Sequence 344, App
99	146.5	19.2	762	US-09-989-731-344	Sequence 344, App
100	146.5	19.2	762	US-09-989-732-344	Sequence 344, App

ALIGNMENTS

RESULT 1
US-10-257-021-53
; Sequence 53, Application US/10257021
; Publication No. US2003021149A1
; GENERAL INFORMATION:
; APPLICANT: Morin, Patrice J.
; APPLICANT: Sherman-Baust, Cheryl A.
; APPLICANT: Pizer, Ellen S.
; APPLICANT: Hough, Colleen D.
; TITLE OF INVENTION: TUMOR MARKERS IN OVARIAN CANCER
; FILE REFERENCE: 14014.0369U2
; CURRENT APPLICATION NUMBER: US/10/257,021
; CURRENT FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: PCT/US01/10947
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/194,336
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-257-021-53

Alignment Scores:
Pred. No.:
Score: 2.21e-82   Length: 399
Percent Similarity: 762.00   Matches: 132
Best Local Similarity: 100.00%   Conservative: 0
Query Match: 100.00%   Mismatches: 0
DB: 15   Indels: 0
                                        Gaps: 0

US-10-613-105-2 (1-132) x US-10-257-021-53 (1-399)

Qy 1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 ATGAAGTCCAGCGGCCTCTTCCCCCTCGTGCTGCTCGCCCTGGGAACTCTGGCACCT 60

Qy 21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysSer 40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 120

Qy 41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 180

Qy 61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 240

Qy 81 ProThrArgArgLysProGlyLysCysProValThrTyrGlnValThrGlyLeuMetLeuAsn 100
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 241 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 300

Qy 101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuCysLysArgLeuCysMetGly 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 360

Qy 121 MetCysGlyLysSerCysValSerProValSerAla 132
        ||||||||||||||||||||||||||||
Db 361 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 396

RESULT 2
US-10-172-118-864
; Sequence 864, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 864
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_003066
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-864

Alignment Scores:
Pred. No.:
Score: 2.21e-82   Length: 399
Percent Similarity: 762.00   Matches: 132
Best Local Similarity: 100.00%   Conservative: 0
Query Match: 100.00%   Mismatches: 0
DB: 15   Indels: 0
                                        Gaps: 0

US-10-172-118-864

US-10-613-105-2 (1-132) x US-10-172-118-864 (1-399)

Qy 1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 ATGAAGTCCAGCGGCCTCTTCCCCCTCGTGCTGCTCGCCCTGGGAACTCTGGCACCT 60

Qy 21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysSer 40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 120

Qy 41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db   121 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCCAGTGTCCAGGGAAG 180

Qy    61 LysArgCysProAspThrCysGlyIleLysCysLeuAspThrCysValAspProAsn  80
Db   181 AAGAGAGTGTGTCCTGACACTTGTGGCATCAAATGCCTGTTGACACCCCAAAC 240

Qy    81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
Db   241 CCAACAAGAGGAGAAGCCTGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 300

Qy   101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
Db   301 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 360

Qy   121 MetCysGlyLysSerCysValSerProValLysAla 132
Db   361 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 396
```

RESULT 3
US-10-342-887-864
; Sequence 864, Application US/10342887
; Publication No. US2004005834OA1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 864
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-864

Alignment Scores:
Pred. No.:   399
Score: 132
Percent Similarity: 2.21e-82   Length: 399
Best Local Similarity: 762.00   Matches: 132
Query Match: 100.00%   Conservative: 0
DB: 100.00%   Mismatches: 0
                          100.00%   Indels: 0
                          16   Gaps: 0

US-10-613-105-2 (1-132) x US-10-342-887-864 (1-399)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
Db   1 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTCTTGGCCCTGGGAACTCTGGCACCT 60

Qy   21 TrpAlaGluGluSerGlyIleValLysSerPheLysAlaGlyValCysProLeuProLysSer 40
Db   61 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCTAAGAAATCT 120

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAsnTrpProGlnCysProGlyLys 60
Db   121 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCCAGTGTCCAGGGAAG 180

Qy   61 LysArgCysProAspThrCysGlyIleLysCysLeuAspThrCysValAspProAsn 80
Db   181 AAGAGAGTGTGTCCTGACACTTGTGGCATCAAATGCCTGTTGACACCCCAAAC 240
```

RESULT 4
US-10-309-290-141
; Sequence 141, Application US/10309290
; Publication No. US2004002324lAl
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Anderson, David W.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Chillakuru, Rajeev A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: Gorman, Linda
; APPLICANT: Gould-Rothberg, Bonnie E.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Jeffers, Michael E.
; APPLICANT: Ji, Weizhen
; APPLICANT: Li, Li
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Miller, Charles E.
; APPLICANT: Murphey, Ryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Peyman, John A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Smithson, Glennda
; APPLICANT: Starling, Gary
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zhong, Haihong
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-502A
; CURRENT APPLICATION NUMBER: US/10/309,290
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: 60/336,600
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,285
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/341,346
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/341,477
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/341,540
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/342,592
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/344,297
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/344,903
; PRIOR FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/373,288
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/380,981
; PRIOR FILING DATE: 2002-05-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 274
; SOFTWARE: CuraSeqlist version 0.1
; SEQ ID NO 141

; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15)..(410)
US-10-309-290-141

Alignment Scores:
Pred. No.:   2.36e-82   Length: 422
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: Indels: 0
DB: Gaps: 16

US-10-613-105-2 (1-132) x US-10-309-290-141 (1-422)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro 20
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   15 ATGAAGTCCAGCGGGCCCTTCTCCCCTTCCCTGGTGCTGCTGCTGGGAACTCTGGCACCT 74

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysGlyValCysProProLysLysSer 40
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGGTGTGTGTCCTCCTAAGAAATCT 134

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   135 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 194

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   195 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTTGATCCTGTTGACACCCCAAAC 254

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   255 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 314

Qy   101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   315 CCCCCAAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 374

Qy   121 MetCysGlyLysSerCysValSerProValLysAla 132
         ||||||||||||||||||||||||||||||||||
Db   375 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 410

RESULT 5
US-10-360-368-1
; Sequence 1, Application US/10360368
; Publication No. US20030220281A1
; GENERAL INFORMATION:
; APPLICANT: Nielson, Lorna I
; APPLICANT: Zhen, Li
; TITLE OF INVENTION: Treatment Of Endometriosis With Antileukoprotease
; FILE REFERENCE: 018002-002020US
; CURRENT APPLICATION NUMBER: US/10/360,368
; CURRENT FILING DATE: 2003-02-06
; PRIOR APPLICATION NUMBER: 09/605,134
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,157
; PRIOR FILING DATE: 1999-07-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
; LENGTH: 565
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (59)..(457)
; OTHER INFORMATION:
US-10-360-368-1

Alignment Scores:
Pred. No.:   3.35e-82   Length: 565
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: Indels: 0
DB: Gaps: 15

US-10-613-105-2 (1-132) x US-10-360-368-1 (1-565)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro 20
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   59 ATGAAGTCCAGCGGGCCCTTCTCCCCTTCCCTGGTGCTGCTGCTGGGAACTCTGGCACCT 118

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysGlyValCysProProLysLysSer 40
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   119 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGTGTGTGTCCTCCTAAGAAATCT 178

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   179 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 238

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   239 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTTGACCCTGTTGACACCCCAAAC 298

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   299 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 358

Qy   101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   359 CCCCCAAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 418

Qy   121 MetCysGlyLysSerCysValSerProValLysAla 132
         ||||||||||||||||||||||||||||||||||
Db   419 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 454

RESULT 6
US-09-964-824A-582
; Sequence 582, Application US/09964824A
; Patent No. US2002010253lA1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatur
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
SEQ ID NO 582
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-582

Alignment Scores:
Pred. No.:   3.55e-82   Length: 594
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: Indels: 0
DB: Gaps: 9

US-10-613-105-2 (1-132) x US-09-964-824A-582 (1-594)

DB: US-10-613-105-2 (1-132) x US-09-954-456-1989 (1-594)

Gaps: 0

US-09-954-456-1989

```
Qy     1  MetLysSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro  20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    19  ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGGCTCTGGGCACCT  78

Qy    21  TrpAlaValGluGlySerGlyLysSerPheLysGlyValCysProProLysSer  40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    79  TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGGTGTGTCTCCTCCTAAGAAATCT  138

Qy    41  AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   139  GCCCAGTGCCTAGATACAAGAAACCTGAGTGCCAGTGACTGGCAGTGTCCAGGGAAG  198

Qy    61  LysArgCysLeuArgProGlnAspThrCysGlyIleLysCysLeuAspThrProAsn  80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   199  AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATACCCCCAAAC  258

Qy    81  ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn  100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   259  CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC  318

Qy   101  ProProAsnPheCysGluMetAspGlyLysLysArgAspLeuLysCysArgLeuLysCysMetGly  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   319  CCCCCCAATTTCTGTGAGATGGATGGCAAGCGTGACTTGAAGTGTCGCATGGGC  378

Qy   121  MetCysGlyLysSerCysValSerProValLysAla  132
          |||||||||||||||||||||||||||||||||||||
Db   379  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  414
```

RESULT 7
US-09-954-456-1989
; Sequence 1989, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; ;   APPLICANT: Young, Paul
; ;   TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; ;   TITLE OF INVENTION: Sets
; ;   FILE REFERENCE: 689290-76
; ;   CURRENT APPLICATION NUMBER: US/09/954,456
; ;   CURRENT FILING DATE: 2001-09-18
; ;   PRIOR APPLICATION NUMBER: US/60/233,617
; ;   PRIOR FILING DATE: 2000-09-18
; ;   PRIOR APPLICATION NUMBER: US/60/234,052
; ;   PRIOR FILING DATE: 2000-09-20
; ;   PRIOR APPLICATION NUMBER: US/60/234,923
; ;   PRIOR FILING DATE: 2000-09-25
; ;   PRIOR APPLICATION NUMBER: US/60/235,134
; ;   PRIOR FILING DATE: 2000-09-25
; ;   PRIOR APPLICATION NUMBER: US/60/235,637
; ;   PRIOR FILING DATE: 2000-09-26
; ;   PRIOR APPLICATION NUMBER: US/60/235,638
; ;   PRIOR FILING DATE: 2000-09-27
; ;   PRIOR APPLICATION NUMBER: US/60/235,711
; ;   PRIOR FILING DATE: 2000-09-27
; ;   PRIOR APPLICATION NUMBER: US/60/235,720
; ;   PRIOR FILING DATE: 2000-09-27
; ;   PRIOR APPLICATION NUMBER: US/60/235,840
; ;   PRIOR FILING DATE: 2000-09-27
; ;   PRIOR APPLICATION NUMBER: US/60/235,863
; ;   PRIOR FILING DATE: 2000-09-27
; ;   NUMBER OF SEQ ID NOS: 2276
; ;   SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1989
; ;   LENGTH: 594
; ;   TYPE: DNA
; ;   ORGANISM: Homo sapiens
; US-09-954-456-1989

Alignment Scores:
Pred. No.:   /
Score: 3.55e-82   Length: 594
                        762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: Indels: 0

RESULT 8
US-09-865-812-1
; Sequence 1, Application US/09865812
; Patent No. US20020115626A1
; GENERAL INFORMATION:
; ;   APPLICANT: Rastelli, Luca
; ;   APPLICANT: Smithson, Glennda
; ;   TITLE OF INVENTION: Method of Detecting inflammatory Lung Disorders
; ;   FILE REFERENCE: 21402-018 US
; ;   CURRENT APPLICATION NUMBER: US/09/865,812
; ;   CURRENT FILING DATE: 2001-05-28
; ;   PRIOR APPLICATION NUMBER: 60/207,104
; ;   PRIOR FILING DATE: 2000-05-25
; ;   NUMBER OF SEQ ID NOS: 5
; ;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; ;   LENGTH: 594
; ;   TYPE: DNA
; ;   ORGANISM: Homo sapiens
; US-09-865-812-1

Alignment Scores:
Pred. No.:   /
Score: 3.55e-82   Length: 594
                        762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: Indels: 0
DB: Gaps: 0

DB: US-10-613-105-2 (1-132) x US-09-865-812-1 (1-594)

```
Qy     1  MetLysSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro  20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    19  ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGGCTCTGGGCACCT  78

Qy    21  TrpAlaValGluGlySerGlyLysSerPheLysGlyValCysProProLysSer  40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    79  TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGGTGTGTCTCCTCCTAAGAAATCT  138

Qy    41  AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

US-10-613-105-2 (1-132) x US-10-097-340-295 (1-594)

```
Qy    1 MetLysSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
Db   19 ATGAAGTCCAGGCCCTTCCCCTTCCTCGTGCTGCTGGCTGGTGCCCTGGGAACTCTGGCACCT  78

Qy   21 TrpAlaValGluGlySerGlyLysPheLysSerPheLysValCysProLysSer  40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGGCTGTGGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCGTCCTCTAAGAAATCT 138

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy   61 LysArgCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  199 AAGAGATGTTGTCCTGACACTTGTGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyValCysLeuMetLeuAsn 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysMetGly 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319 CCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
         ||||||||||||||||||||||||||||||||||||
Db  379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414
```

Alignment Scores:
Pred. No.:
3.55e-82   Length: 594

```
Db  139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy   61 LysArgCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  199 AAGAGATGTTGTCCTGACACTTGTGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyValCysLeuMetLeuAsn 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysMetGly 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319 CCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
         ||||||||||||||||||||||||||||||||||||
Db  379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414
```

Alignment Scores:
Pred. No.:
Score: 3.55e-82   Length: 594
Percent Similarity: 762.00   Matches: 132
Best Local Similarity: 100.00%   Conservative: 0
Query Match: 100.00%   Mismatches: 0
DB: 14   Indels: 0
                                            Gaps: 0

```
Score:                  762.00      Matches:         132
Percent Similarity:     100.00%     Conservative:    0
Best Local Similarity:  100.00%     Mismatches:      0
Query Match:            100.00%     Indels:          0
DB:                                 Gaps:            0

US-10-613-105-2 (1-132) x US-10-205-823-374 (1-594)

Qy    1 MetLysSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 ATGAAGTCCAGCGGGCCCTCTCCCCTGTGCTTGCCCTGGGAACTCTGGCACCT 78

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProLysSer 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGGCTGTGGAAGGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 138

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnSerProGlyLys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy   61 LysArgCysCysProAspThrCysGlyIleLeuCysGluLeuAspProValAspThrProAsn 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCTGTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259 CCAACAAGGAGGAAGCCTGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319 CCCCCAAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414
```

RESULT 11
US-10-177-293-429
; Sequence 429, Application US/10177293
; Publication No. US2003012412BA1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 429
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-293-429

Alignment Scores:
```
Pred. No.:                                     Length:          594
Score:                  3.55e-82               Matches:         132
Percent Similarity:     100.00%                Conservative:    0
Best Local Similarity:  100.00%                Mismatches:      0
Query Match:            100.00%                Indels:          0
DB:                                            Gaps:            0
```

US-10-613-105-2 (1-132) x US-10-177-293-429 (1-594)

```
Qy    1 MetLysSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 ATGAAGTCCAGCGGGCCCTCTCCCCTGTGCTTGCCCTGGGAACTCTGGCACCT 78

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProLysSer 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGGCTGTGGAAGGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 138

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnSerProGlyLys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy   61 LysArgCysCysProAspThrCysGlyIleLeuCysGluLeuAspProValAspThrProAsn 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCTGTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259 CCAACAAGGAGGAAGCCTGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319 CCCCCAAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414
```

RESULT 12
US-10-172-118-863
; Sequence 863, Application US/10172118
; Publication No. US2003022437A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 863
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Homo sapiens

```
;    PUBLICATION INFORMATION:
;    DATABASE ACCESSION NUMBER: NM 003064
;    DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-863

Alignment Scores:
Pred. No.:
Score:                        3.55e-82
                              762.00
Percent Similarity:           100.00%
Best Local Similarity:        100.00%
Query Match:
DB:                           15

US-10-613-105-2 (1-132) x US-10-172-118-863  (1-594)

Qy    1 MetLysSerGlyLeuPheProPheLeuValLeuAlaAlaLeuGlyThrLeuAlaPro 20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 ATGAAGTCCAGGGGCCTCTTCCCCTTCCCTTCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCT 78

Qy   21 TrpAlaValGluGlyLysSerGlyLysSerPheLysSerAlaGlyValCysProProLysLysSer 40
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGGCTGTGGAAGGCTCTGGAAAGTCTTCAAAGCTGGAGTCTGTCCTTCCTAAGAAATCT 138

Qy   41 AlaGlnCysLeuAlaArgTyrArgLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy   61 LysArgCysProAspThrCysGlyIleLysCysLeuAspThrProAsn 80
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGGATCCTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysProValThrTyrGlyValCysLeuMetLeuAsn 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy  121 MetCysGlyLysLysSerCysValSerProValLysAla 132
        ||||||||||||||||||||||||||||||||||||||||
Db  379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414
```

RESULT 13
US-10-305-720-1271
; Sequence 1271, Application US/10305720
; Publication No. US2004010136A1
; GENERAL INFORMATION:
;   APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
;   TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expressio
;   FILE REFERENCE: PA-0002-1 CON
;   CURRENT APPLICATION NUMBER: US/10/305,720
;   CURRENT FILING DATE: 2002-11-26
;   PRIOR APPLICATION NUMBER: 09/016,434
;   PRIOR FILING DATE: 1998-01-30
;   NUMBER OF SEQ ID NOS: 1490
;   SOFTWARE: PERL Program
;   SEQ ID NO 1271
;   LENGTH: 594
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: GenBank ID No. US2004010136A1 g28638
US-10-305-720-1271

Alignment Scores:
Pred. No.:
Score: 3.55e-82
                              762.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match:
DB: 16

```
US-10-613-105-2 (1-132) x US-10-305-720-1271 (1-594)

Qy    1 MetLysSerGlyLeuPheProPheLeuValLeuAlaAlaLeuGlyThrLeuAlaPro 20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 ATGAAGTCCAGGGGCCTCTTCCCCTTCCCTTCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCT 78

Qy   21 TrpAlaValGluGlyLysSerGlyLysSerPheLysSerAlaGlyValCysProProLysLysSer 40
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGGCTGTGGAAGGCTCTGGAAAGTCTTCAAAGCTGGAGTCTGTCCTTCCTAAGAAATCT 138

Qy   41 AlaGlnCysLeuAlaArgTyrArgLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy   61 LysArgCysProAspThrCysGlyIleLysCysLeuAspThrProAsn 80
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysProValThrTyrGlyValCysLeuMetLeuAsn 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy  121 MetCysGlyLysLysSerCysValSerProValLysAla 132
        ||||||||||||||||||||||||||||||||||||||||
Db  379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414
```

RESULT 14
US-10-309-290-143
; Sequence 143, Application US/10309290
; Publication No. US2004002324A1
; GENERAL INFORMATION:
;   APPLICANT: Alsobrook II, John P.
;   APPLICANT: Anderson, David W.
;   APPLICANT: Boldog, Ferenc L.
;   APPLICANT: Burgess, Catherine E.
;   APPLICANT: Chillakuru, Rajeev A.
;   APPLICANT: Edinger, Shlomit R.
;   APPLICANT: Gerlach, Valerie L.
;   APPLICANT: Gorman, Linda
;   APPLICANT: Gould-Rothberg, Bonnie E.
;   APPLICANT: Guo, Xiaojia
;   APPLICANT: Jeffers, Michael E.
;   APPLICANT: Ji, Weizhen
;   APPLICANT: Li, Li
;   APPLICANT: Malyankar, Uriel M.
;   APPLICANT: Miller, Charles E.
;   APPLICANT: Murphey, Ryan
;   APPLICANT: Patturajan, Meera
;   APPLICANT: Peyman, John A.
;   APPLICANT: Rastelli, Luca
;   APPLICANT: Rieger, Daniel K.
;   APPLICANT: Shenoy, Suresh G.
;   APPLICANT: Smithson, Glennda
;   APPLICANT: Starling, Gary
;   APPLICANT: Taupier, Raymond J.
;   APPLICANT: Voss, Edward Z.
;   APPLICANT: Zhong, Haihong
;   APPLICANT: Zhong, Mei
;   TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
;   FILE REFERENCE: 21402-502A
;   CURRENT APPLICATION NUMBER: US/10/309,290
;   CURRENT FILING DATE: 2002-12-02
;   PRIOR APPLICATION NUMBER: 60/336,600
;   PRIOR FILING DATE: 2001-12-05
;   PRIOR APPLICATION NUMBER: 60/338,285
;   PRIOR FILING DATE: 2001-12-07
;   PRIOR APPLICATION NUMBER: 60/341,346
;   PRIOR FILING DATE: 2001-12-12

; PRIOR APPLICATION NUMBER: 60/341,477
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/341,540
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/342,592
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/344,297
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/344,903
; PRIOR FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/373,288
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/380,981
; PRIOR FILING DATE: 2002-05-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 274
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 143
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(414)
US-10-309-290-143

Alignment Scores:
Pred. No.:   3.55e-82   Length: 594
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 16   Indels: 0
                                        Gaps: 0

US-10-613-105-2 (1-132) x US-10-309-290-143 (1-594)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuValLeuAlaLeuGlyThrLeuAlaPro 20
Db   19 ATGAAGTCCAGCGGCCGCTCTTCCCCTTCCGTGCTGCTTGCCCTGGGAACTCTGGCACCT 78

Qy   21 TrpAlaValGluGluGlySerPheLysSerAlaGlyLysValCysProProLysProLysSer 40
Db   79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAAGTCCTTAAGAAATCT 138

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLys 60
Db 139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTGACTGGCAGTGTCCAGGGAAG 198

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysGlyLeuAspProValAspThrProAsn 80
Db 199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysValLeuMetLeuAsn 100
Db 259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTGATGCTTAAC 318

Qy 101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
Db 319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy 121 MetCysGlyLysSerCysValSerProValLysAla 132
Db 379 ATGTGTGGGAAATCCTGCGTTTCCCGTGAAAGCT 414

RESULT 15
US-10-240-425-1529
; Sequence 1529, Application US/10240425
; Publication No. US2004003352A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1529
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US2004003352A1 X04470
US-10-240-425-1529

Alignment Scores:
Pred. No.:   3.55e-82   Length: 594
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 16   Indels: 0
DB: Gaps: 0

US-10-613-105-2 (1-132) x US-10-240-425-1529 (1-594)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuValLeuAlaLeuGlyThrLeuAlaPro 20
Db   19 ATGAAGTCCAGCGGCCGCTCTTCCCCTTCCGTGCTGCTTGCCCTGGGAACTCTGGCACCT 78

Qy   21 TrpAlaValGluGluGlySerPheLysSerAlaGlyLysValCysProProLysProGlyLys 40
Db   79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAAGTCGTCCAGGGAAG 138

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
Db 139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTGACTGGCAGTGTCCAGGGAAG 198

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysGlyLeuAspProValAspThrProAsn 80
Db 199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysValLeuMetLeuAsn 100
Db 259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTGATGCTTAAC 318

Qy 101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
Db 319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy 121 MetCysGlyLysSerCysValSerProValLysAla 132
Db 379 ATGTGTGGGAAATCCTGCGTTTCCCGTGAAAGCT 414

RESULT 16
US-10-342-887-863
; Sequence 863, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 863
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-863
; LENGTH: 594

Alignment Scores:
Pred. No.:
Score: 3.55e-82
Percent Similarity: 762.00
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 100.00%

Length: 594
Matches: 132
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 16

US-10-613-105-2 (1-132) x US-10-342-887-863 (1-594)

```
Qy    1 MetLysSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGTGCTTGCTTGCCCTGGGAACTCTGCACCT  78

Qy   21 TrpAlaValGluGlyLysSerGlyLysSerPheLysAlaGlyLysValCysProLysLysSer  40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 138

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319 CCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
        ||||||||||||||||||||||||||||||||||||
Db  379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414
```

RESULT 17
US-10-613-105-1
; Sequence 1, Application US/10613105
; Publication No. US20040115690A1
; GENERAL INFORMATION:
; APPLICANT: Rastelli, Luca
; APPLICANT: Smithson, Glennda
; TITLE OF INVENTION: Method of Detecting inflammatory Lung Disorders
; FILE REFERENCE: 21402-018 DIV
; CURRENT APPLICATION NUMBER: US/10/613,105
; CURRENT FILING DATE: 2003-07-02
; PRIOR APPLICATION NUMBER: 60/207,104
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 09/865,812
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 594

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-613-105-1

Alignment Scores:
Pred. No.:
Score: 3.55e-82
Percent Similarity: 762.00
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 100.00%

Length: 594
Matches: 132
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 17

US-10-613-105-2 (1-132) x US-10-613-105-1 (1-594)

```
Qy    1 MetLysSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGTGCTTGCTTGCCCTGGGAACTCTGCACCT  78

Qy   21 TrpAlaValGluGlyLysSerGlyLysSerPheLysAlaGlyLysValCysProLysLysSer  40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 138

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy  101 ProProAsnPheCysGluGlnSerAspTrpGlnCysLysArgAspLeuLysCysCysMetGly 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319 CCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
        ||||||||||||||||||||||||||||||||||||
Db  379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414
```

RESULT 18
US-10-322-696-125
; Sequence 125, Application US/10322696
; Publication No. US20040166490A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Malandro, Marc
; TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
; FILE REFERENCE: 52945201200
; CURRENT APPLICATION NUMBER: US/10/322,696
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 125
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-696-125

Alignment Scores:
Pred. No.:
Score: 3.55e-82
Percent Similarity: 762.00
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 100.00%

Length: 594
Matches: 132
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 17

US-10-613-105-2 (1-132) x US-10-322-696-125 (1-594)

```
Qy    1 MetLysSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Alignment Scores:
Pred. No.:   3.58e-82   Length: 598
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 0

US-10-613-105-2 (1-132) x US-10-177-293-431 (1-598)

```
Qy      1 MetLysSerGlyLeuProPheLeuValLeuAlaLeuGlyThrLeuAlaPro  20
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23 ATGAAGTCCAGGGGCCTCTTCCCCTTCCTGGTCGTGCTTGCCCTGGGAACTCTGGCACCT  82

Qy     21 TrpAlaValGluGlySerPheLysSerAlaGlyValCysProProLysSer  40
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     83 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGCCCTCCTCCTAAGAAATCT  142

Qy     41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    143 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG  202

Qy     61 LysArgCysCysProAspThrCysGlyIleIleLysCysLeuAspProValAspThrAsn  80
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 AAGAGATGTTGTCCTGACACTTGTGGCATCAAAATGCCTGGATCCTGTTGACACCCCAAAC  262

Qy     81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyValCysLeuMetLeuAsn  100
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTGATGCTTAAC  322

Qy    101 ProProAsnPheCysGluMetCysAspGlyGlnCysLysArgAspLeuLysCysMetGly  120
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    323 CCCCCAATTTCTGTGAGATGGCAAGCCAGGTGGCAAGCGTGACTTGAAGTGTTGCATGGC  382

Qy    121 MetCysGlyLysSerCysValSerProValLysAla  132
            ||||||||||||||||||||||||||||||||||||
Db    383 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  418
```

RESULT 20
US-10-173-999-153
; Sequence 153, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 153
;   LENGTH: 598
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-173-999-153

Alignment Scores:
Pred. No.:   3.58e-82   Length: 598
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0

(second column / lower panel)

```
Db     19 ATGAAGTCCAGGGGCCTCTTCCCCTTCCTGGTCGTGCTTGCCCTGGGAACTCTGGCACCT  78

Qy     21 TrpAlaValGluGlySerPheLysSerAlaGlyValCysProProLysSer  40
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGCCCTCCTCCTAAGAAATCT  138

Qy     41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG  198

Qy     61 LysArgCysCysProAspThrCysGlyIleIleLysCysLeuAspProValAspThrAsn  80
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAAATGCCTGGATCCTGTTGACACCCCAAAC  258

Qy     81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyValCysLeuMetLeuAsn  100
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTGATGCTTAAC  318

Qy    101 ProProAsnPheCysGluMetCysAspGlyGlnCysLysArgAspLeuLysCysMetGly  120
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    319 CCCCCAATTTCTGTGAGATGGCAAGCCAGGTGGCAAGCGTGACTTGAAGTGTTGCATGGC  378

Qy    121 MetCysGlyLysSerCysValSerProValLysAla  132
            ||||||||||||||||||||||||||||||||||||
Db    379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  414
```

RESULT 19
US-10-177-293-431
; Sequence 431, Application US/10/177,293
; Publication No. US2003012412BA1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 431
;   LENGTH: 598
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-177-293-431

```
Pred. No.:                7.28e-82    Length:         1084
Score:                    762.00     Matches:        132
Percent Similarity:       100.00%    Conservative:   0
Best Local Similarity:    100.00%    Mismatches:     0
Query Match:              100.00%    Indels:         0
DB:                       10         Gaps:           0

US-10-613-105-2 (1-132) x US-09-814-353-20587 (1-1084)

Qy     1  MetLysSerGlyLeuPhePheLeuValLeuAlaLeuGlyThrLeuAlaLeuAlaPro  20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    69  ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTTGCCCTGGGAACTCTGGCACCT  128

Qy    21  TrpAlaValGluGlySerGlyLysPheLysSerProLysAlaGlyValCysProProLysLysSer  40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   129  TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT  188

Qy    41  AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnSerProGlyLys  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   189  GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG  248

Qy    61  LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   249  AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC  308

Qy    81  ProThrArgArgLysProGlyLysCysProValThrTyrGlyValGlnCysLeuMetLeuAsn  100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   309  CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTGATGCTTAAC  368

Qy   101  ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   369  CCCCCAAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC  428

Qy   121  MetCysGlyLysSerValSerProValLysAla  132
             ||||||||||||||||||||||||||||
Db   429  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  464
```

RESULT 22
US-10-025-514-15
; Sequence 15, Application US/10025514
; Publication No. US20030073217A1
; GENERAL INFORMATION:
; APPLICANT: Philip J. BARR
; APPLICANT: Helen GIBSON
; APPLICANT: Philip PEMBERTON
; TITLE OF INVENTION: MULTIFUNCTIONAL PROTEASE INHIBITORS AND
; TITLE OF INVENTION: THEIR USE IN TREATMENT OF DISEASE
; FILE REFERENCE: 36829200200
; CURRENT APPLICATION NUMBER: US/10/025,514
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: U.S. 60/256,699
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: U.S. 60/331,966
; PRIOR FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-025-514-15

Alignment Scores:
```
Pred. No.:                7.65e-67    Length:         1525
Score:                    638.50     Matches:        112
Percent Similarity:       87.97%     Conservative:   5
Best Local Similarity:    84.21%     Mismatches:     9
Query Match:              83.79%     Indels:         7
DB:                       14         Gaps:           1

US-10-613-105-2 (1-132) x US-10-025-514-15 (1-1525)

Qy     7  PheProPheLeuValLeu----------------LeuAlaLeuGlyThrLeuAla  19
```

```
Best Local Similarity:    100.00%    Mismatches:     0
Query Match:              100.00%    Indels:         0
DB:                       16         Gaps:           0

US-10-613-105-2 (1-132) x US-10-173-999-153 (1-598)

Qy     1  MetLysSerGlyLeuPhePheLeuValLeuAlaLeuGlyThrLeuAlaPro  20
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db    23  ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTTGCCCTGGGAACTCTGGCACCT  82

Qy    21  TrpAlaValGluGlySerGlyLysPheLysSerProLysAlaGlyValCysProProLysLysSer  40
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db    83  TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT  142

Qy    41  AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnSerProGlyLys  60
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   143  GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG  202

Qy    61  LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   203  AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC  262

Qy    81  ProThrArgArgLysProGlyLysCysProValThrTyrGlyValGlnCysLeuMetLeuAsn  100
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   263  CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTGATGCTTAAC  322

Qy   101  ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly  120
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   323  CCCCCAAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC  382

Qy   121  MetCysGlyLysSerValSerProValLysAla  132
             ||||||||||||||||||||||||||||
Db   383  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  418
```

RESULT 21
US-09-814-353-20587
; Sequence 20587, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20587
; LENGTH: 1084
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 1083, 1084
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-20587

Alignment Scores:

```
Db  1119 TTCGTTTTCTGATGATCGAGCAGCAGACACTAAAAGCCCATTGTTTATGGGTAAGGTTGTC 1178
              |||      |||||||::::::                          |||      :::|||
Qy    20 ProTrpAlaValGluGlySerGlyPheLysAlaGlyValCysProProLysLys         39
Db  1179 AACCCAACTCAGAGAGATGTCCGGAAAGTCTTTCAAGGCCGGTGTTTGTCCCCAAAGAAG 1238
Qy    40 SerAlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpCysGlnCysProGly 59
Db  1239 TCCGCTCAATGTTTGAGATACAAGAGCCAGAATGTCAATCCGACTGGCAATGTCCAGGT 1298
Qy    60 LysLysArgCysMetAspThrCysGlyIleLeuLysCysLeuAspProValAspThrPro   79
Db  1299 AAGAAGAGATGTTGTCCAGACACTTGTGGTATCAAGTGTCTAGACCCAGTTGACACCCCA 1358
Qy    80 AsnProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeu   99
Db  1359 AACCCAACTAGAGAAGAAAGCCAGTAAGTGTCCAGTTACTTACGGTCAATGTTTGATGTTG 1418
Qy   100 AsnProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMet  119
Db  1419 AACCCACCAAACTTCTGTGAAATGGACGGTCAATGTAAGAGAGACTTGAAGTGTTGTATG 1478
Qy   120 GlyMetCysGlyLysSerCysValSerProValLysAla                       132
Db  1479 GGTATGTGTGGTAAGTCCTGTGTTTCCCCAGTCAAGGCC                      1517
```

RESULT 23
US-10-025-514-3
; Sequence 3, Application US/10025514
; Publication No. US2003007317A1
; GENERAL INFORMATION:
; APPLICANT: Philip J. BARR
; APPLICANT: Helen GIBSON
; APPLICANT: Philip PEMBERTON
; TITLE OF INVENTION: MULTIFUNCTIONAL PROTEASE INHIBITORS AND
; TITLE OF INVENTION: THEIR USE IN TREATMENT OF DISEASE
; FILE REFERENCE: 36829200200
; CURRENT APPLICATION NUMBER: US/10/025,514
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: U.S. 60/256,699
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: U.S. 60/331,966
; PRIOR FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-025-514-3

Alignment Scores:

Pred. No.:	3.14e-67	Length:	321
Score:	635.00	Matches:	107
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	83.33%	Indels:	0
DB:	14	Gaps:	0

US-10-613-105-2 (1-132) x US-10-025-514-3 (1-321)

```
Qy    26 SerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeuArg     45
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TCTGGAAAGTCTTTCAAGGCCGGTGTTTGTCCACCAAAGAGTGCCCTCAATGTTTGAGA   60
Qy    46 TyrLysLysProGluCysGlnSerAspTrpCysGlnSerProGlyLysArgCysPro     65
Db    61 TACAAGAAGCCAGAATGTCAATCCGACTGGCAATGTCCAGGTAAGAGAGATGTTGTCCA  120
Qy    66 AspThrCysGlyIleLeuLysCysLeuAspProValAspProAsnProThrArgArgLys   85
Db   121 GACACTTGTGGTATCAAGTGTCTAGAGACCCAGTTGACACCCCAACCCAACTAGAGAAAG  180
```

RESULT 24
US-10-309-290-145
; Sequence 145, Application US/10309290
; Publication No. US20040023241A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Anderson, David W.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Chillakuru, Rajeev A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: Gorman, Linda
; APPLICANT: Gould-Rothberg, Bonnie E.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Jeffers, Michael E.
; APPLICANT: Ji, Weizhen
; APPLICANT: Li, Li
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Miller, Charles E.
; APPLICANT: Murphey, Ryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Peyman, John A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Smithson, Glennda
; APPLICANT: Starling, Gary
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zhong, Haihong
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-502A
; CURRENT APPLICATION NUMBER: US/10/309,290
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: 60/336,600
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,285
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/341,346
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/341,477
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/341,540
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/342,592
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/344,297
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/344,903
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/373,288
; PRIOR FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/380,981
; PRIOR FILING DATE: 2002-04-17
; PRIOR FILING DATE: 2002-05-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 274
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 145

```
Qy    86 ProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsnProProAsnPheCys  105
Db   181 CCAGGTAAGTGTCCAGTTACTTACGGTCAATGTTTGATGTTGAACCCACCAAACTTCTGT  240
Qy   106 GluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMetCysGlyLysSer  125
Db   241 GAAATGGACGGTCAATGTAAGAGAGACTTGAAGTGTTGTATGGGTATGTGTGGTAAGTCC  300
Qy   126 CysValSerProValLysAla                                         132
Db   301 TGTGTTTCCCCAGTCAAGGCC                                         321
```

```
;   LENGTH: 321
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(321)
US-10-309-290-145

Alignment Scores:
Pred. No.:                          321
Score:               3.14e-67       107
Percent Similarity:  635.00         0
Best Local Similarity: 100.00%      0
Query Match:         100.00%        0
DB:                  83.33%
                     16

US-10-613-105-2 (1-132) x US-10-309-290-145  (1-321)

Qy     26  SerGlyLysSerPheLysAlaGlyValCysProProLysLysProLysSerAlaGlnCysLeuArg  45
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  TCTGGAAAGTCTTCAAAGCTGGAGTCTGTCTGCCCCAGAAATCTGCCCAGTGCCTTAGA  60

Qy     46  TyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysPro  65
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  TACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAGAAGAGATGTTGTCCT  120

Qy     66  AspThrCysGlyIleLysCysLeuAspProValAspProAsnProThrArgArgLys  85
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAACCCAACGAGGAGAAG  180

Qy     86  ProGlyLysCysProValThrTyrGlyIleCysLeuMetLeuAsnProProAsnPheCys  105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  CCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAACCCCCCAATTTCTGT  240

Qy    106  GluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMetCysGlyLysSer  125
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  GAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGCATGTGTGGGAAATCC  300

Qy    126  CysValSerProValLysAla  132
           ||||||||||||||||
Db    301  TGCGTTTCCCGTGAAAGCT  321
```

```
RESULT 25
US-10-025-514-7
;   Sequence 7, Application US/10025514
;   Publication No. US20030073217A1
;   GENERAL INFORMATION:
;   APPLICANT: Phillip J. BARR
;   APPLICANT: Helen GIBSON
;   APPLICANT: Philip PEMBERTON
;   TITLE OF INVENTION: MULTIFUNCTIONAL PROTEASE INHIBITORS AND
;   TITLE OF INVENTION: THEIR USE IN TREATMENT OF DISEASE
;   FILE REFERENCE: 36829200200
;   CURRENT APPLICATION NUMBER: US/10/025,514
;   CURRENT FILING DATE: 2002-04-03
;   PRIOR APPLICATION NUMBER: U.S. 60/256,699
;   PRIOR FILING DATE: 2000-12-18
;   PRIOR APPLICATION NUMBER: U.S. 60/331,966
;   PRIOR FILING DATE: 2001-11-20
;   NUMBER OF SEQ ID NOS: 33
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 7
;   LENGTH: 1525
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-025-514-7

Alignment Scores:
Pred. No.:                          1525
Score:               2.02e-66       107
Percent Similarity:  635.00         0
Best Local Similarity: 100.00%      0
Query Match:         100.00%
                     83.33%
```

```
DB:                                 14              Gaps:           0

US-10-613-105-2 (1-132) x US-10-025-514-7  (1-1525)

Qy     26  SerGlyLysSerPheLysAlaGlyValCysProProLysLysProLysSerAlaGlnCysLeuArg  45
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     12  TCTGGAAAGTCTTCAAAGCTGGAGTCTGTCTGCCCCAGAAGAAGAGAGTCCGCTCAATGTTTGAGA  71

Qy     46  TyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysPro  65
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     72  TACAAGAAGCCAGAATGTCAATCCGACTGGCAATGTCCAGGTAAGAGAAGAGATGTTGTCCA  131

Qy     66  AspThrCysGlyIleLysCysLeuAspProValAspProAsnProThrArgArgLys  85
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    132  GACACTTGTGGTATCAAGTGCTTAGACCCAGTTGACACCCCAAACCCAACTAGAAGAAAG  191

Qy     86  ProGlyLysCysProValThrTyrGlyIleCysLeuMetLeuAsnProProAsnPheCys  105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    192  CCAGGTAAGTGCCCAGTTACTTACGGTCAATGTTTGATGTTGAACCCACCAAACTTCTGT  251

Qy    106  GluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMetCysGlyLysSer  125
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    252  GAAATGGACGGTCAATGTAAGAGAGACTTGAAGTGTTGTATGGGTATGTGTGGTAAGTCC  311

Qy    126  CysValSerProValLysAla  132
           ||||||||||||||||
Db    312  TGTGTTTCCCCAGTCAAGGCC  332
```

```
RESULT 26
US-10-322-696-124
;   Sequence 124, Application US/10322696
;   Publication No. US2004016690A1
;   GENERAL INFORMATION:
;   APPLICANT: Morris, David W.
;   APPLICANT: Malandro, Marc
;   TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
;   FILE REFERENCE: 52945200200
;   CURRENT APPLICATION NUMBER: US/10/322,696
;   CURRENT FILING DATE: 2003-10-17
;   NUMBER OF SEQ ID NOS: 186
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 124
;   LENGTH: 22324
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-322-696-124

Alignment Scores:
Pred. No.:                          22324
Score:               2.6e-53        103
Percent Similarity:  537.50         2
Best Local Similarity: 42.51%       1
Query Match:         41.70%         141
DB:                  70.54%         1
                     17

US-10-613-105-2 (1-132) x US-10-322-696-124  (1-22324)

Qy     26  SerGlyLysSerPheLysAlaGlyValCysProProLysLysProLysSerAlaGlnCysLeuArg  45
           ||||| ||:::||||||||||||||||||||||||||||||||||||||||||||||::||||
Db  10819  TCTTCAAAAGCCTTCAAAGCTGGAGTCTGTCTCCTCCTAAGAAATCTGCCCAGTGCCTTAGA  10878

Qy     46  TyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysPro  65
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10879  TACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAGAAGAGATGTTGTCCT  10938

Qy     66  AspThrCysGlyIleLysCysLeuAspProValAspProAsnProThr---------  82
           ||||||||||||||||||||||||||||||||||::::
Db  10939  GACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAACCCAAGTAAGCAGTC  10998

Qy     82  ------------------------------------------------------------  82
Db  10999  GGGGAACTGGGTAGAGAGAGTGAGCCTGGGACACAGCATTAGAGGATGGAACTGGGT  11058
```

Qy   82 -- 82
Db   11059 GATGGGTCCTGCCAGCCTCCTTGTCAATCCGTCAGTGAGTCACACTGCCCTAAGCAGGA 11118

Qy   82 -- 82
Db   11119 AGGTAGCCAGCAGCTGGTGAAGCAGCGCGGCCATTAGATAGCCAGGTAGTTGGAAGCCTCC 11178

Qy   82 -- 82
Db   11179 CACCTAGTCAGCACTGGGTCGCTGCCACCTGCGATCAATGGGGGCTCGAAGTTCTAGGA 11238

Qy   82 -- 82
Db   11239 GAGCCAGGTGCTATGTTGGGGGCCGCCCTTAGGGAGGAAGGTGGTGTGATAGAGGTGGGG 11298

Qy   82 -- 82
Db   11299 AGGGGATGATCCCCCGCTGAAGCTGGAGCAGGGGGCTCACTCTAAAAAGTGGGGATGGG 11358

Qy   83 ---------ArgArg 84
Db   11359 AGGGGTTGTATAAAGTACAAGCGCCTCTGACCGTAGCCTCACCTCTCACCCAGCAAGGAGG 11418

Qy   85 LysProGlyLysCysProValThrTyrGlyLysGlnCysLeuMetLeuAsnProProAsnPhe 104
Db   11419 AAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAACCCCCCAATTTC 11478

Qy   105 CysGluMetAspGlyValGlnCysLysArgAspLeuLysCysMetGlyMetCysGlyLys 124
Db   11479 TGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGCATGTGTGGGAAA 11538

Qy   125 SerCysValSerProValLys 131
Db   11539 TCCTGCGTTTCCCCTGTGAAA 11559

RESULT 27
US-10-322-696-122
; Sequence 122, Application US/10322696
; Publication No. US2004016490A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Malandro, Marc
; TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
; FILE REFERENCE: 52945200I200
; CURRENT APPLICATION NUMBER: US/10/322,696
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 122
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-322-696-122

Alignment Scores:
Pred. No.:
Score: 5.49e-46   Length: 1079
Percent Similarity: 72.73%   Matches: 79
Best Local Similarity: 59.85%   Conservative: 17
Query Match: 60.83%   Mismatches: 33
                                            Indels: 3
                                            Gaps: 2
DB: 17

US-10-613-105-2 (1-132) x US-10-322-696-122 (1-1079)

Qy   1 MetLysSerSerGlyLeuPhePheLeuValLeuAlaLeuGlyThrLeuAlaPro 20
Db   440 ATGAAGTCCTCGGGNCTCTTTACCTTTCACCGTGCTCCTTGCCTCTGGGATCCTGGCACCC 499

Qy   21 TrpAlaValGluGlySerGlyLys-----SerPheLysAlaGlyValCysProProLys 38
Db   500 TGGACTGTGGAAGGA--GGCAAAAATGATGCTATCAAAATGCGGAGCCCTGCCCCTGCTAA 556

Qy   39 LysSerAlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysPro 58
Db   557 AAGCCTGCCCAGTGCCTTAAGCTTGAGAAGCCACCACATGCGTACTGGGAGTGCCCG 616

Qy   59 GlyLysGlyArgCysProAspThrCysGlyIleLysCysLysAspProValAspThr 78
Db   617 GGAAAGCAGAGTGTGTGCCAAGATGCTTGCGGTTCCAAGTGCGTGAATCCTGTTCCCATT 676

Qy   79 ProAsnProThrArgArgLysProGlyLysCysProValThrTyrGlyLysGlnCysLeuMet 98
Db   677 CGCAAACCAGTGTGGGAGGAAGCCTGGGAGGTGCGTCAAAACTCAGGCAAGATGTATGATG 736

Qy   99 LeuAsnProProAsnPheCysGluMetAspGlyValGlnCysLysArgAspLeuLysCys 118
Db   737 CTTAACCCTCCCAATGTCTGCCAGAGGGACGGGCAGTGTGACGGCAAATACAAGTGCTGT 796

Qy   119 MetGlyMetCysGlyLysSerCysValSerProVal 130
Db   797 GAGGGTATATGTGGAAAGTCTGCCTGCCCCCGATG 832

RESULT 28
US-10-305-720-1024
; Sequence 1024, Application US/10305720
; Publication No. US2004010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1024
; LENGTH: 325
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US2004010136A1 939088
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) ... (325)
; OTHER INFORMATION: a, t, c, g, or other
US-10-305-720-1024

Alignment Scores:
Pred. No.:
Score: 6.01e-46   Length: 325
Percent Similarity: 89.81%   Matches: 96
Best Local Similarity: 88.89%   Conservative: 1
Query Match: 60.10%   Mismatches: 10
DB: 16   Indels: 4
                                            Gaps: 0

US-10-613-105-2 (1-132) x US-10-305-720-1024 (1-325)

Qy   1 MetLysSerSerGlyLeuPhePheLeuValLeuAlaLeuGlyThrLeuAlaPro 20
Db   1 ATGAAGTCCAGCGGNCTCTTTNCCCTTNCCCTTCTGGTGCTGCCCCTGGGAACTCTGGCACCT 60

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCys-ProProLysSe 40
Db   61 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCAAGGAAATC 120

Qy   40 rAlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLy 60
Db   121 TGCCCAGTGCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAA 180

Qy   60 sLysArgCysProAspThrCysGlyIleLysCysLysAspProValAspThrProAs 80
Db   181 GAAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCTGATCCTNTGACAC-CCAAA 239

Qy   80 nProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAs 100
                ||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   240 CCCAACAAGAGGAGGAAGCT-GGGAAGTGNCCAGTAGTTGNCAATGTTGATGTTAAN 298

Qy   100 nProProAsnPheCysGluMet 107
                ||||||||||||||
Db   299 CCCCCCAA--TTTTGTGAGATG 318

RESULT 29
US-10-125-159-37
; Sequence 37, Application US/10125159
; Publication No. US20030166023A1
; GENERAL INFORMATION:
; APPLICANT: Iartchouk, Natalia
; APPLICANT: Ayers, Mark D.
; APPLICANT: Brown, Jeffrey L.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION AND THERAPY OF
; TITLE OF INVENTION: HUMAN CANCERS
; FILE REFERENCE: MRI-029
; CURRENT APPLICATION NUMBER: US/10/125,159
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: 60/295,031
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/284,773
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/284,764
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-125-159-37

Alignment Scores:
Pred. No.:   3.94e-33   Length: 422
Score: 352.50   Matches: 82
Percent Similarity: 67.44%   Conservative: 5
Best Local Similarity: 63.57%   Mismatches: 40
Query Match: 46.26%   Indels: 5
DB: 15   Gaps: 0

US-10-613-105-2 (1-132) x US-10-125-159-37 (1-422)

Qy   6 LeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaLeuProTrpAlaValGluGly 25
               |||||||||||||||||||||||||||||||||||
Db   27 CTTTTTCTTTCTGGTCGTGAAAAGAAAGGGAACTTCTGGCACCTTGGGCTGAGGAGGC 86

Qy   26 SerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeu-Ar 45
               ||||||||| :|||||| ||||||| ||
Db   87 TCTGGAAAGAACTTCAAGGGTGGAGCCCGCCCTCTAACAAAAGGGCCAGTGCCTGGAG 146

Qy   45 gTyrLys-LysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysP 65
               |||| |||||| ||| ||||||||||||
Db   147 ATACAACCAAACCTGGGGGGCAGAGGGACTGCGCCCC-CCCAGGAGGAAGAGAGCCCC-C 204

Qy   65 roAspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArgArgL 85
               ||||| |:||||
Db   205 CTGACACTCGGGGGAATCAAATGCCTGAGGGTGAAGACACCTT-AAACCCCACCAGGAGGA 263

Qy   85 ysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsnProProAsnPheC 105
               ||||||||||||||||||||||||||||||
Db   264 AGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTGAAGAAAAAGGGGGGGGGTTAT 323

Qy   105 ysGluMetAspGlyLysGlnCysArgAspLeuLysCysCysSerGlyLeuCysGlyLysS 125
               |||||| ||||||||| |:::
Db   324 GAGAAAAGGATGGCCAGAGAACAAGCGTGACTTGAAGCGGTGCCTCCTCATGTGTGGGAGAG 383

Qy   125 erCysValSerProValLysAla 132
               :|||
Db   384 GCTGCGAAACCCCTGGCAAAGCT 406

RESULT 30
US-10-116-802-431
; Sequence 431, Application US/10116802
; Publication No. US2003006515A1
; GENERAL INFORMATION:
; APPLICANT: Amy Lasek
; TITLE OF INVENTION: GENES EXPRESSED IN LUNG CANCER
; FILE REFERENCE: PA-0045 US
; CURRENT APPLICATION NUMBER: US/10/116,802
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/281,593
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 519
; SOFTWARE: PERL Program
; SEQ ID NO 431
; LENGTH: 805
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 1159769.1
US-10-116-802-431

Alignment Scores:
Pred. No.:   3.57e-31   Length: 805
Score: 339.00   Matches: 104
Percent Similarity: 72.19%   Conservative: 5
Best Local Similarity: 68.87%   Mismatches: 21
Query Match: 44.49%   Indels: 21
DB: 14   Gaps: 3

US-10-613-105-2 (1-132) x US-10-116-802-431 (1-805)

Qy   1 MetLysSerSerGlyLeu-PheProPheLeuValLeuLeuAla-LeuGlyThrLeu-Ala 19
               |||||||||||||| ||||||||||||||||||| |||||||| ||
Db   124 ATGAAGTCCAGGCCCTCTTTCCCCTGTGCTTCCTCTGTGCTCCTTGCCCTTGGGGAACTCTGGACA 183

Qy   20 ProTrpAla-ValGluGly-SerGlyLysSerPheLys--AlaGlyValCys-ProProL 38
               ||||| |||||| :|||||| ||| |||||||
Db   184 CCTTGGGCTTGTGGAAGGACTCCTGGAAAGTCCCTCAATAGCTGGAGTCTGTCCCTCCAA 243

Qy   38 ysLysSerAlaGlnCysLeu--ArgTyrLysLysProGlu-CysGlnSerAspTrp-Gl 56
               :||||||||| ||| :::|||||||||||| |||||||||||||||
Db   244 GGAAATCTGCCCAGATGCCTTAGCATACAAGGAAACCTGATGTGCCAGAGTGACTGGGCA 303

Qy   56 nCysProGlyLysLysArgCysCys-ProAspThrCysGlyIle-LysCysLeu-AspPr 75
               |||||||||||||| ||||| |||||||||||||||| |||||||||
Db   304 GTGTCCAGGGGAGACGAGATGTTGTCCCTGACACTTGGGCATCCAAATGCCTGGGATCC 363

Qy   75 oValAspThr-ProAsnProThrArgArgLysPro-GlyLysCysProVal-ThrTyrG1 94
               ||||||| |||||||||||||||||||| |||||||||||||| |||||||
Db   364 TGTTGACACCCCAAACCCACAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGG 423

Qy   94 yGlnCysLeu-MetLeuAsnProProAsnPheCys-----GluMetAspGlyLysGlnCysL 112
               |||||||| |||||||||||||||||||| ||||||||||
Db   424 CCAATGTTTGAAGGCTTAACCCCCCCAATTTTTGTGAGATGATGGTCCAGTGCAA 483

Qy   112 ysArgAspLeuLysCysCysMetGlyMet---CysGlyLysSerCysValSerProVal 130
               |||||||||||| |||:::
Db   484 GCTGTGACTTGAGTGTTGCATGGGCATGATGTGTGGAAATCCTGTCGATTTCCCCTG 542

RESULT 31
US-10-076-816-57
; Sequence 57, Application US/10076816
; Publication No. US2003005624A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Ning
; APPLICANT: Rodriguez, Raymond
; APPLICANT: Hagie, Frank E.
; TITLE OF INVENTION: Feed Additive Compositions and Methods
; FILE REFERENCE: 50665-8021.US00
; CURRENT APPLICATION NUMBER: US/10/076,816

```
Db   2522 AGCCCTAAGACATCGCCCTGTGAATCCTTCTCACAGGTGTGGAGGAAGCCTGGAGG 2581
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     89 CysProValThrTyrGlyGlnCysLeuMetLeuAsnProProAsnPheCysGluMetAsp 108
            |||      |||  |||::::|||||||:::|||||:::|||||||
Db   2582 TGCGTCAAAACTACGGAAGATGTATGATGTCTTAACCCTCCAATGTCTGCCAGAGGAC 2641
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    109 GlyGlnCysLysArgAspLeuLysCysMetGlyMetCysValLysSerCysValSer 128
            ||||||||||||||||||||||||||||||||::::|||||||||::::
Db   2642 GGGCAGTGTGACGGCAAATACAGTGCTGTGAGGGTATATGTGGGAAAGTCTGCCTGCCC 2701
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    129 ProVal 130
            ||:::
Db   2702 CCGATG 2707
```

RESULT 32
US-09-814-353-2826

```
; Sequence 2826, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
;   APPLICANT: Lee, John
;   APPLICANT: Thompson, Pamela
;   APPLICANT: Lillie, James
;   TITLE OF INVENTION:  NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;   TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;   TITLE OF INVENTION:  THERAPY OF OVARIAN CANCER
;   FILE REFERENCE: MRI-006B
;   CURRENT APPLICATION NUMBER: US/09/814,353
;   CURRENT FILING DATE: 2001-03-21
;   PRIOR APPLICATION NUMBER: US 60/191,031
;   PRIOR FILING DATE: 2000-03-21
;   PRIOR APPLICATION NUMBER: US 60/207,124
;   PRIOR FILING DATE: 2000-05-25
;   PRIOR APPLICATION NUMBER: US 60/211,940
;   PRIOR FILING DATE: 2000-06-15
;   PRIOR APPLICATION NUMBER: US 60/216,820
;   PRIOR FILING DATE: 2000-07-07
;   PRIOR APPLICATION NUMBER: US 60/220,661
;   PRIOR FILING DATE: 2000-07-25
;   PRIOR APPLICATION NUMBER: US 60/257,672
;   PRIOR FILING DATE: 2000-12-21
;   NUMBER OF SEQ ID NOS: 22037
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2826
;   LENGTH: 313
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 14, 20, 24, 50, 51, 53, 58, 61, 73, 95, 117, 134, 202, 222,
;   LOCATION: 265, 267, 270, 272, 287, 292, 301, 308
;   OTHER INFORMATION: n = A,T,C or G
US-09-814-353-2826
```

Alignment Scores:
```
Pred. No.:                                5.43e-25        Length:          313
Score:                                   283.50          Matches:          55
Percent Similarity:                       79.17%         Conservative:      2
Best Local Similarity:                    76.39%         Mismatches:       14
Query Match:                              37.20%         Indels:            1
DB:                                       10             Gaps:              1
```

US-10-613-105-2 (1-132) x US-09-814-353-2826 (1-313)

```
Qy     1 MetLysSerSerGlyLeuProPheProLeuValLeuAlaLeuGlyThrLeuAlaPro 20
            :::                |||||||||||||||||||||
Db    67 TTGAACNCCAGCGCGCCTCTTCCCCCTTTCCNGGGGCTGCTTGCCCTGGGAACNCTGGCACCT 126
            |||||||||||||||||||||||||||||||||||||||||||||
Qy    21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer 40
            |||||||||||||||||||||||||||
Db   127 TGGGCTGNGGAAGGCCTCGAAAGTCCTTCAAAGCTCGAGTCTGTCCTCCTAAGAAATCT 186
            ||||||||||||||
Qy    41 AlaGlnCysLeuArgTyrLysLysProGluCysLysGlnSerAspTrpCysPro---Gly 59
            |||||||||||||||
Db      ArgLysProGlyLys 88
            |||:::
```

```
;   CURRENT FILING DATE:   2002-02-14
;   PRIOR APPLICATION NUMBER: US 60/269,188
;   PRIOR FILING DATE: 2001-02-14
;   PRIOR APPLICATION NUMBER: US 09/847,232
;   PRIOR FILING DATE: 2001-05-02
;   PRIOR APPLICATION NUMBER: US 60/266,929
;   PRIOR FILING DATE: 2001-02-06
;   PRIOR APPLICATION NUMBER: US 60/201,182
;   PRIOR FILING DATE: 2000-05-02
;   NUMBER OF SEQ ID NOS: 60
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 57
;   LENGTH: 4090
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(4090)
;   OTHER INFORMATION: n = A,T,C or G
US-10-076-816-57
```

Alignment Scores:
```
Pred. No.:                                7.69e-24        Length:         4090
Score:                                   285.00          Matches:          57
Percent Similarity:                       28.24%         Conservative:     17
Best Local Similarity:                    21.76%         Mismatches:       28
Query Match:                              37.40%         Indels:          160
DB:                                                      Gaps:             14
```

US-10-613-105-2 (1-132) x US-10-076-816-57 (1-4090)

```
Qy    29 SerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeuArgTyrLysLys 48
            :::  |||  :::||||||||||||||||||||||  :::|||
Db  1922 GCTATCAAAATCGGAGCCCTGCCCTGCTAAAAAGCCTGCCCAGTGCCTTAAGCTTGAGAAG 1981
Qy    49 ProGluCysGlnSerAspTrpTrpGlnCysProGlyLysProLysArgCysProAspThrCys 68
            |||:::|||||||:::||||||  |||||||:::|||
Db  1982 CCACAATGCCCATACTGACTGGAGTGCCCCGGGAAAGCAAGAGGTGCTGCCAAGATGCTTGC 2041
Qy    69 GlyIleLysCysLeuAspProValAspThrProAsnProThrArg---------- 83
            |||:::|||  :::|||||||
Db  2042 GGTTCCAAGTGCGTGAATCCTGTTCCCATTCGCAAACCAGTGAGCCAGCAGAGAAAACAG 2101
Qy    83 ---------------------------------------------- 83
Db  2102 AGAGGGGACAAATCAACGTCAAGAGCCACCCAGAGTGAATGAAGAGTCCCTCCCAGGCATC 2161
Qy    83 ---------------------------------------------- 83
Db  2162 CTTGTTTCTAAGAGCTGTTCGGGTCCAACATGTCAGTAGTCAAGGCCTCTTGCTGCCATT 2221
Qy    83 ---------------------------------------------- 83
Db  2222 GGCAGGAAGAGCCATTTGCCTGCAGTGCCACTCGTGTGAGGATGATGGCTTACAACCAT 2281
Qy    83 ---------------------------------------------- 83
Db  2282 TCGGCCCTTTCGGGGCTGGAGAGCTCGTCCCCTGCATTGGCTGGAGGCTAGATGGTAGCG 2341
Qy    83 ---------------------------------------------- 83
Db  2342 AGGAAGCAGGTCCTCCTTCAGGAGGCGTGTGTTCGGGAGGAGGAGGTGACAAAGGTGATGGGTGG 2401
Qy    83 ---------------------------------------------- 83
Db  2402 ATGGGGACGAGACTTCCTCGCTCAGGAAGGAGAGATGGTTGGGTTAAAGACACATGGAGATGTGTT 2461
Qy    83 ---------------------------------------------- 83
Db  2462 GCCTGAGCCCTAAGACATCGCCCTGTGAATCCTTTCTCAACAGGGAGAATGTGTTGCCTG 2521
Qy    84 ---------------------------------------------- 84
```

RESULT 33
US-09-814-353-9156
; Sequence 9156, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: PatentIn version 4.0
; SEQ ID NO 9156
; LENGTH: 313
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 14, 20, 24, 50, 51, 53, 58, 61, 73, 95, 117, 134, 202, 222,
; LOCATION: 265, 267, 270, 272, 287, 292, 301, 308
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-9156

Alignment Scores:
Pred. No.:   5.43e-25   Length: 313
Score: 283.50   Matches: 55
Percent Similarity: 79.17%   Conservative: 2
Best Local Similarity: 76.39%   Mismatches: 14
Query Match: 37.20%   Indels: 1
DB: 10   Gaps: 1

US-10-613-105-2 (1-132) x US-09-814-353-9156 (1-313)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
           ::: ||||||||| ||||||::|||||||||||||||||||
Db   67 TTGAACNNCAGCGGCCTCTTCCCCTTCCnGGGCCTGCTTGCCCTGGGAACNCTGGCACCT 126

Qy   21 TrpAlaValGluGlySerPheLysSerPheLysAlaGlyValCysProProLysLysSer 40
           |||||||||:||||||||||| |||||||||||||||||| ||||||||||||||||||
Db   127 TGGGCTGNGGAGGGCTCTGAAAGTCCTTCAAAGCTNGAAGCTGGAGTGTGTCCTCCTAAGAATCT 186

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysPro---Gly 59
           ||||||||||||:|| || |||| ||| |||||| ||||
Db   187 GCCCAGTGCCTTAGANACAAGAAACCTGAGTGCCANAGTGACTGGCAGGGCCAAGGGAA 246

Qy   60 LysLysArgCysProAspThrCysGlyIleLys 71
           |||||:::| ||||||
Db   247 AAAAAAAAAGTTGCCCTNANCCNTNGGGGAAAAA 282

RESULT 34
US-10-322-696-121
; Sequence 121, Application US/10322696
; Publication No. US20040166490A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Malandro, Marc
; TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
; FILE REFERENCE: 52945200120O
; CURRENT APPLICATION NUMBER: US/10/322,696
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn for Windows Version 4.0
; SEQ ID NO 121
; LENGTH: 38692
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-322-696-121

Alignment Scores:
Pred. No.:   8.95e-22   Length: 38692
Score: 277.50   Matches: 57
Percent Similarity: 30.33%   Conservative: 17
Best Local Similarity: 23.36%   Mismatches: 28
Query Match: 36.42%   Indels: 143
DB: Gaps: 1

US-10-613-105-2 (1-132) x US-10-322-696-121 (1-38692)

Qy   29 SerPheLysAlaGlyValCysProProLysLysSerAlaGlnCysLeuArgTyrLysLys 48
            ::: ||| ||| ||||||||||| |||||: ::|||
Db 27153 GCTATCAAAATCGGAGCCTGCCCTGCTAAAAAGCCTGCCCAGTGCCTTAAGCTTGAGAAG 27212

Qy   49 ProGluGlnSerAspTrpGlnCysProGlyLysProGlyLysArgCysProAspThrCys 68
            ||| ||||||||||||:|||||||:||:||:|||||
Db 27213 CCACAATGCCGTACTGACTGGGAGTGCCCGGGAAAGCAGAGAGTGCTGCCAAGATGCTTGC 27272

Qy   69 GlyIleLysCysLeuAspProValAspThrProAsnPro 81
            ||| |||||| ||:||
Db 27273 GGTTCCAAGTGCGTGAATCCTGTTCCCATTGCAAACCAGTGAGCAGCAGCAGAGAAAACA 27332

Qy   81 --- 81

Db 27333 GAGAGGGGACAATCAACGTCAAGAGCCACCACCAGAGTGAATGAAGAGTCCCTCCCAGGCAT 27392

Qy   81 --- 81

Db 27393 CCTTGTTCTAAGAGCTGTTGGGTCCAACATGTCAGTAGTCAAGGCCTCTTGCTGCCAT 27452

Qy   81 --- 81

Db 27453 TGGCAGGAAGAGCCATTTGCCTGCAGTTGCCACTCTGTGAGGATGATGGCTTTACACCA 27512

Qy   81 --- 81

Db 27513 TTCGGGCCCTTTCGGGGCTGAGAGAGCTGGTCCCGTGCCCTGCATTGGCTGGCGAGGCTAGC 27572

Qy   81 --- 81

Db 27573 GAGGAAGCAGGTCCTCTTCAGGAGGCGTGTTGGGAGGAGGTGACAAAGGTGATGGGTG 27632

Qy   81 --- 81

Db 27633 GATGGGACGAACTTCCTCGCTCAGGAAGGAGATGTTGGGTTAAAGACATGGAGAATGTGT 27692

Qy   82 ---ThrArgLysPro 86
           --- |||
Db 27693 TGCCTGACCCTAAGACATCGGCCCCTGTGAATCCTTTCTCAACAGTGTGG-AGGAGGCCT 27751

Qy   87 GlyLysCysProValThrTyrGlyGlnThrLeuMetLeuAsnProProAsnPheCysGlu 106
            :::::|| | ||||| ||| || || :||:|:||:|:
Db 27752 GGGAGGTGCGTCAAAACTCAGGCAAGATGATGTCTTAACCCTCCCAATGTCTGCCAG 27811

Qy   107 MetAspGlyGlnIleCysLysArgAspLeuLysCysGlyMetGlyGluLysSerCys 126
           |||||||||||||| |||::|||

```
Db      27812  AGGGACGGGCAGTGTGACGGCCAAATACAAGTGCTGTGAGGGTATATGTGGGAAAGTCTGC  27871

Qy        127  ValSerProVal  130
                 ::: |||:::
Db      27872  CTGCCCCGATG  27883

RESULT 35
US-09-964-824A-495/c
; Sequence 495, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
;   APPLICANT: Horrigan, Stephen
;   TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
;   TITLE OF INVENTION: Sets
;   FILE REFERENCE: 689290-73
;   CURRENT APPLICATION NUMBER: US/09/964,824A
;   CURRENT FILING DATE: 2001-09-27
;   PRIOR APPLICATION NUMBER: US/60/236,033
;   PRIOR FILING DATE: 2000-09-28
;   PRIOR APPLICATION NUMBER: US/60/236,032
;   PRIOR FILING DATE: 2000-09-28
;   PRIOR APPLICATION NUMBER: US/60/236,028
;   PRIOR FILING DATE: 2000-09-28
;   NUMBER OF SEQ ID NOS: 583
;   SOFTWARE: PatentIn version 3.0
; SEQ ID NO 495
;   LENGTH: 411
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(411)
;   OTHER INFORMATION: n=a,t,g or c
US-09-964-824A-495

Alignment Scores:
Pred. No.:              1.66e-22        Length:         411
Score:                  264.00          Matches:        53
Percent Similarity:     74.67%          Conservative:   3
Best Local Similarity:  70.67%          Mismatches:     17
Query Match:            34.65%          Indels:         3
DB:                     9               Gaps:           1

US-10-613-105-2 (1-132) x US-09-964-824A-495 (1-411)

Qy      60  LysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrPro  79
               |||:::||||       ||::|||    |||||
Db     406  AAAAGAAATNTTTGCCCNGAAACTTGGNGCATNAANTGCCG-GGNCCCCTGTNNACCCCC  348

Qy      80  ---AsnProThrArgArgLysPro-GlyLysCysProValThrTyrGlyGlnCysLeuMe  98
               |||||                  |||       |||||
Db     347  CNAANCCCAACAAGGNGNAACCCTTGGNAAGTNCCNAGTGANTTNTGGCCAATGTTNGAT  288

Qy      98  tLeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCy  118
               |||||||||||||||      |||||||||||
Db     287  GCTNAACCCCCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTG  228

Qy     118  sMetGlyMetCysGlyLysSerCysValSerProValLysAla  132
               |||||||||||||||
Db     227  CATGGGCATGTGTGGGAAAATCCTGCGTTTCCCCTGTGAAAGCT  185

RESULT 36
US-09-954-456-1987/c
; Sequence 1987, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
;   APPLICANT: Young, Paul
;   TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
;   TITLE OF INVENTION: Sets
;   FILE REFERENCE: 689290-76
;   CURRENT APPLICATION NUMBER: US/09/954,456
;   CURRENT FILING DATE: 2001-09-18
;   PRIOR APPLICATION NUMBER: US/60/233,617
```

```
;   PRIOR FILING DATE: 2000-09-18
;   PRIOR APPLICATION NUMBER: US/60/234,052
;   PRIOR FILING DATE: 2000-09-20
;   PRIOR APPLICATION NUMBER: US/60/234,923
;   PRIOR FILING DATE: 2000-09-25
;   PRIOR APPLICATION NUMBER: US/60/235,134
;   PRIOR FILING DATE: 2000-09-25
;   PRIOR APPLICATION NUMBER: US/60/235,637
;   PRIOR FILING DATE: 2000-09-26
;   PRIOR APPLICATION NUMBER: US/60/235,638
;   PRIOR FILING DATE: 2000-09-26
;   PRIOR APPLICATION NUMBER: US/60/235,711
;   PRIOR FILING DATE: 2000-09-27
;   PRIOR APPLICATION NUMBER: US/60/235,720
;   PRIOR FILING DATE: 2000-09-27
;   PRIOR APPLICATION NUMBER: US/60/235,840
;   PRIOR FILING DATE: 2000-09-27
;   PRIOR APPLICATION NUMBER: US/60/235,863
;   PRIOR FILING DATE: 2000-09-27
;   NUMBER OF SEQ ID NOS: 2276
;   SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1987
;   LENGTH: 411
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: n=a,t,g or c
US-09-954-456-1987

Alignment Scores:
Pred. No.:              1.66e-22        Length:         411
Score:                  264.00          Matches:        53
Percent Similarity:     74.67%          Conservative:   3
Best Local Similarity:  70.67%          Mismatches:     17
Query Match:            34.65%          Indels:         3
DB:                     9               Gaps:           1

US-10-613-105-2 (1-132) x US-09-954-456-1987 (1-411)

Qy      60  LysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrPro  79
               |||:::||||       ||::|||    |||||
Db     406  AAAAGAAATNTTTGCCCNGAAACTTGGNGCATNAANTGCCG-GGNCCCCTGTNNACCCCC  348

Qy      80  ---AsnProThrArgArgLysPro-GlyLysCysProValThrTyrGlyGlnCysLeume  98
               |||||                  |||       |||||
Db     347  CNAANCCCAACAAGGNGNAACCCTTGGNAAGTNCCNAGTGANTTNTGGCCAATGTTNGAT  288

Qy      98  tLeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCy  118
               |||||||||||||||      |||||||||||
Db     287  GCTNAACCCCCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTG  228

Qy     118  sMetGlyMetCysGlyLysSerCysValSerProValLysAla  132
               |||||||||||||||
Db     227  CATGGGCATGTGTGGGAAAATCCTGCGTTTCCCCTGTGAAAGCT  185

RESULT 37
US-09-967-768A-3/c
; Sequence 3, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
;   APPLICANT: Augustus, Meena
;   TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
;   TITLE OF INVENTION: Sets
;   FILE REFERENCE: 689290-72
;   CURRENT APPLICATION NUMBER: US/09/967,768A
;   CURRENT FILING DATE: 2001-09-28
;   PRIOR APPLICATION NUMBER: US/60/236,109
;   PRIOR FILING DATE: 2000-09-28
;   PRIOR APPLICATION NUMBER: US/60/236,034
;   PRIOR FILING DATE: 2000-09-28
;   PRIOR APPLICATION NUMBER: US/60/236,111
;   PRIOR FILING DATE: 2000-09-28
```

```
;     NUMBER OF SEQ ID NOS: 325
;     SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
;     LENGTH: 411
;     TYPE: DNA
;     ORGANISM: Homo sapiens
;     FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (1)...(411)
;     OTHER INFORMATION: n=a,t,g or c
US-09-967-768A-3

Alignment Scores:
Pred. No.:               1.66e-22
Score:                   264.00
Percent Similarity:      74.67%
Best Local Similarity:   70.67%
Query Match:             34.65%
DB:

US-10-613-105-2 (1-132) x US-09-967-768A-3 (1-411)

Qy    60 LysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrPro 79
         |||:::|                              |||:
Db   406 AAAAGAAATNTTGCCNGAAACTTGGNGCATNAANTGCCG-GGNCCCTGTNNACCCCC 348

Qy    80 ---AsnProThrArgArgLysPro-GlyLysCysProValThrTyrGlyGlnCysLeuMe 98
                                                           |||||
Db   347 CNAANCCAACAAGGNGNAACCCTTGGNAAGTNCCNAGTGANTTNTGGCCAATGTTNGAT 288

Qy    98 tLeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysLeuLysAspLeuLysCysCy 118
         |||||||||||||||||||||||||||||||||||||
Db   287 GCTNAACCCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTG 228

Qy   118 sMetGlyMetCysGlyLysSerCysValSerProValLysAla 132
         |||||||||||||||||||||||||||||
Db   227 CATGGGCATGTGTGGGAAATCCTGCGTTCCCCTGTGAAAGCT 185
```

```
RESULT 38
US-09-960-706-16/c
; Sequence 16, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
;     APPLICANT: Munger, William E.
;     TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplas
;     TITLE OF INVENTION: Gene Expression Profiles
;     FILE REFERENCE: 44921-5029-01US
;     CURRENT APPLICATION NUMBER: US/09/960,706
;     CURRENT FILING DATE: 2001-09-24
;     PRIOR APPLICATION NUMBER: 60/223,323
;     PRIOR FILING DATE: 2000-08-07
;     PRIOR APPLICATION NUMBER: 09/873,319
;     PRIOR FILING DATE: 2001-06-05
;     NUMBER OF SEQ ID NOS: 1124
;     SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
;     LENGTH: 411
;     TYPE: DNA
;     ORGANISM: Homo sapiens
;     FEATURE:
;     OTHER INFORMATION: Genbank Accession No. US20030134280A1 AA026641
;     NAME/KEY: unsure
;     LOCATION: (1)..(411)
;     OTHER INFORMATION: n = a or c or g or t
US-09-960-706-16

Alignment Scores:
Pred. No.:               1.66e-22
Score:                   264.00
Percent Similarity:      74.67%
Best Local Similarity:   70.67%
Query Match:             34.65%
DB:

Length:                  411
Matches:                 53
Conservative:            3
Mismatches:              17
Indels:                  3
Gaps:                    1

Length:                  411
Matches:                 53
Conservative:            3
Mismatches:              17
Indels:                  3
Gaps:                    10
```

```
US-10-613-105-2 (1-132) x US-09-960-706-16 (1-411)

Qy    60 LysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrPro 79
         |||:::|                              |||:
Db   406 AAAAGAAATNTTGCCNGAAACTTGGNGCATNAANTGCCG-GGNCCCTGTNNACCCCC 348

Qy    80 ---AsnProThrArgArgLysPro-GlyLysCysProValThrTyrGlyGlnCysLeuMe 98
                                                           |||||
Db   347 CNAANCCAACAAGGNGNAACCCTTGGNAAGTNCCNAGTGANTTNTGGCCAATGTTNGAT 288

Qy    98 tLeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysLeuLysAspLeuLysCysCy 118
         |||||||||||||||||||||||||||||||||||||
Db   287 GCTNAACCCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTG 228

Qy   118 sMetGlyMetCysGlyLysSerCysValSerProValLysAla 132
         |||||||||||||||||||||||||||||
Db   227 CATGGGCATGTGTGGGAAATCCTGCGTTCCCCTGTGAAAGCT 185
```

```
RESULT 39
US-09-873-319-9/c
; Sequence 9, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
;     APPLICANT: Munger, William E.
;     APPLICANT: Kulkarni, Prakash
;     APPLICANT: Getzenberg, Robert H.
;     APPLICANT: Waga, Iwao
;     APPLICANT: Yamamoto, Jun
;     TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
;     TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
;     FILE REFERENCE: 44921-5029-US
;     CURRENT APPLICATION NUMBER: US/09/873,319A
;     CURRENT FILING DATE: 2001-06-05
;     EARLIER APPLICATION NUMBER: US 60/223,323
;     EARLIER FILING DATE: 2000-08-07
;     NUMBER OF SEQ ID NOS: 755
;     SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;     LENGTH: 411
;     TYPE: DNA
;     ORGANISM: Homo sapiens
;     FEATURE:
;     OTHER INFORMATION: Genbank Accession No. US20030134324A1 AA026641
;     FEATURE:
;     NAME/KEY: unsure
;     LOCATION: (1)..(411)
;     OTHER INFORMATION: n = a or c or g or t
US-09-873-319-9

Alignment Scores:
Pred. No.:               1.66e-22
Score:                   264.00
Percent Similarity:      74.67%
Best Local Similarity:   70.67%
Query Match:             34.65%
DB:

Length:                  411
Matches:                 53
Conservative:            3
Mismatches:              17
Indels:                  3
Gaps:                    1
```

```
US-10-613-105-2 (1-132) x US-09-873-319-9 (1-411)

Qy    60 LysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrPro 79
         |||:::|                              |||:
Db   406 AAAAGAAATNTTGCCNGAAACTTGGNGCATNAANTGCCG-GGNCCCTGTNNACCCCC 348

Qy    80 ---AsnProThrArgArgLysPro-GlyLysCysProValThrTyrGlyGlnCysLeuMe 98
                                                           |||||
Db   347 CNAANCCAACAAGGNGNAACCCTTGGNAAGTNCCNAGTGANTTNTGGCCAATGTTNGAT 288

Qy    98 tLeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysLeuLysAspLeuLysCysCy 118
         |||||||||||||||||||||||||||||||||||||
Db   287 GCTNAACCCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTG 228

Qy   118 sMetGlyMetCysGlyLysSerCysValSerProValLysAla 132
```

```
Db    227 CATGGGCATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 185
          ||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 40
US-10-430-201-2236
; Sequence 2236, Application US2004016679A1
; Publication No. US2004016679A1
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; TITLE OF INVENTION: Method for Predicting Gene Potential and Cell Commitment
; FILE REFERENCE: 40716 (IP-010)
; CURRENT APPLICATION NUMBER: US/10/430,201
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US 60/370,114
; PRIOR FILING DATE: 2002-05-03
; NUMBER OF SEQ ID NOS: 4879
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2236
; LENGTH: 248
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (72)..(86)
; OTHER INFORMATION: n is a, c, g, or t
US-10-430-201-2236

Alignment Scores:
Score: 5.48e-22
Pred. No.:
Percent Similarity: 68.75%
Best Local Similarity: 55.00%
Query Match: 33.79%
DB:

Length: 248
Matches: 44
Conservative: 11
Mismatches: 22
Indels: 3
Gaps: 2

US-10-613-105-2 (1-132) x US-10-430-201-2236 (1-248)

```
Qy     15 LeuGlyThrLeuAlaProTrpAlaValGluGlySerGlyLys------SerPheLysAla 32
          ||||||||||||||||||||||||||||||            :: :||
Db      3 CTGGGGATCCTGGCACCCTGGACTGTGGAAGGA---GGCAAAAATGATGCTATCAAAATC 59

Qy     33 GlyValCysProProLysLysSerAlaGlnCysLeuArgTyrLysLysProGluCysGln 52
          ||||||       ||||||||||||:::  :::||||||||::||||||||||
Db     60 GGAGCCTGCCCTNNNNNNNNNNNNNCAGTGCCTTAAGCTTGAGAAGCCACAATGCCGT 119

Qy     53 SerAspTrpGlnCysLysArgCysCysProAspThrCysGlyIleLysCys 72
          ::||||||:::||||||||||:||||||    ||| ||||||||||| ||||||
Db    120 ACTGACTGGGAGTGCCCGGGAAAGCAGAGGTGCTGCCAAGATGCTTGCGGTTCCAAGTGC 179

Qy     73 LeuAspProValAspThrProAsnProThrArgArgLysProGlyLysCysProValThr 92
          :::::||||||                      ||||||||||:||||
Db    180 GTGAATCCTGTTCCCATTCGCAAACCAGTGTGGGAGGAGCCTGGGAGGTGCGTCAAAACT 239
```

US-10-430-201-2237

RESULT 41
US-10-430-201-2237
; Sequence 2237, Application US2004016679A1
; Publication No. US2004016679A1
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; TITLE OF INVENTION: Method for Predicting Gene Potential and Cell Commitment
; FILE REFERENCE: 40716 (IP-010)
; CURRENT APPLICATION NUMBER: US/10/430,201
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US 60/370,114
; PRIOR FILING DATE: 2002-05-03
; NUMBER OF SEQ ID NOS: 4879
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2237
; LENGTH: 248
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (72)..(86)
; OTHER INFORMATION: n is a, c, g, or t
US-10-430-201-2237

Alignment Scores:
Pred. No.:   5.48e-22
Score: 257.50
Percent Similarity: 68.75%
Best Local Similarity: 55.00%
Query Match: 33.79%
DB:

Length: 248
Matches: 44
Conservative: 11
Mismatches: 22
Indels: 3
Gaps: 2

US-10-613-105-2 (1-132) x US-10-430-201-2237 (1-248)

```
Qy     15 LeuGlyThrLeuAlaProTrpAlaValGluGlySerGlyLys------SerPheLysAla 32
          ||||||||||||||||||||||||||||||            :: :||
Db      3 CTGGGGATCCTGGCACCCTGGACTGTGGAAGGA---GGCAAAAATGATGCTATCAAAATC 59

Qy     33 GlyValCysProProLysLysSerAlaGlnCysLeuArgTyrLysLysProGluCysGln 52
          ||||||       ||||||||||||:::  :::||||||||::||||||||||
Db     60 GGAGCCTGCCCTNNNNNNNNNNNNNCAGTGCCTTAAGCTTGAGAAGCCACAATGCCGT 119

Qy     53 SerAspTrpGlnCysLysArgCysCysProAspThrCysGlyIleLysCys 72
          ::||||||:::||||||||||:||||||    ||| ||||||||||| ||||||
Db    120 ACTGACTGGGAGTGCCCGGGAAAGCAGAGGTGCTGCCAAGATGCTTGCGGTTCCAAGTGC 179

Qy     73 LeuAspProValAspThrProAsnProThrArgArgLysProGlyLysCysProValThr 92
          :::::||||||                      ||||||||||:||||
Db    180 GTGAATCCTGTTCCCATTCGCAAACCAGTGTGGGAGGAGCCTGGGAGGTGCGTCAAAACT 239
```

RESULT 42
US-09-814-353-15540
; Sequence 15540, Application US20030165831A1
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15540
; LENGTH: 188
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-15540

Alignment Scores:
Pred. No.:   2.18e-20
Score: 243.00
Percent Similarity: 95.92%
Best Local Similarity: 93.88%
Query Match: 31.89%
DB:

Length: 188
Matches: 46
Conservative: 1
Mismatches: 2
Indels: 0
Gaps: 0

US-10-613-105-2 (1-132) x US-09-814-353-15540 (1-188)

```
; LOCATION: (72)..(86)
; OTHER INFORMATION: n is a, c, g, or t
US-10-430-201-2237
```

```
Qy    1 MetLysSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro 20
           |::     ||||||||||||||||||||||||||||||||
Db   41 ATTAACTCCAGCGGCCTCTTCCCTTCCCTTGCGTGCTGGCCTGCCTGGCACCT 100

Qy   21 TrpAlaValGluGlyGlySerGlyLysPheLysSerPheLysAlaGlyValCysProProLysSer 40
           ||||||||||||||||||||||||||||||||||||
Db  101 TGGGCTGGGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 160

Qy   41 AlaGlnCysLeuArgTyrLysLysPro 49
           ||||||||||||||||
Db  161 GCCCAGTGCCTTAGATACAAGAAACCT 187

RESULT 43
US-10-125-159-43
; Sequence 43, Application US/10125159
; Publication No. US20030166023A1
; GENERAL INFORMATION:
;  APPLICANT: Iartchouk, Natalia
;  APPLICANT: Ayers, Mark D.
;  APPLICANT: Brown, Jeffrey L.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR THE
;  TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION AND THERAPY OF
;  TITLE OF INVENTION: HUMAN CANCERS
;  FILE REFERENCE: MRI-029
;  CURRENT APPLICATION NUMBER: US/10/125,159
;  CURRENT FILING DATE: 2002-04-18
;  PRIOR APPLICATION NUMBER: 60/295,031
;  PRIOR FILING DATE: 2002-05-31
;  PRIOR APPLICATION NUMBER: 60/284,773
;  PRIOR FILING DATE: 2001-04-18
;  PRIOR APPLICATION NUMBER: 60/284,764
;  PRIOR FILING DATE: 2001-04-18
;  NUMBER OF SEQ ID NOS: 88
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 43
;  LENGTH: 518
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-10-125-159-43

Alignment Scores:
Pred. No.:              1.34e-18        Length:          518
Score:                  232.50          Matches:         43
Percent Similarity:     83.02%          Conservative:    1
Best Local Similarity:  81.13%          Mismatches:      6
Query Match:            30.51%          Indels:          3
DB:                     15              Gaps:            1

US-10-613-105-2 (1-132) x US-10-125-159-43 (1-518)

Qy   80 AsnProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeu 99
           |||||        |||   ||||
Db  303 AACATATCAAACACTCCCGGGT-------GTACGCGGGGGCCAATGTTTGATGCTT 353

Qy  100 AsnProProAsnPheCysGluMetGlyGlnCysLysArgLysAspLeuLysCysCysMet 119
           ||||||||||||||||||   ::||||||||||
Db  354 AACCCCCCAATTTCTGTGAGATGGAGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATG 413

Qy  120 GlyMetCysGlyLysSerCysValSerProValLysAla 132
           ||||||||||||||||
Db  414 GGCATGTGTGGGAAAATCCTGCGTTCCCCTGTGAAAGCT 452

RESULT 44
US-10-125-159-41
; Sequence 41, Application US/10125159
; Publication No. US20030166023A1
; GENERAL INFORMATION:
;  APPLICANT: Iartchouk, Natalia
;  APPLICANT: Ayers, Mark D.
;  APPLICANT: Brown, Jeffrey L.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR THE
;  TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION AND THERAPY OF
;  TITLE OF INVENTION: HUMAN CANCERS
;  FILE REFERENCE: MRI-029
;  CURRENT APPLICATION NUMBER: US/10/125,159
;  CURRENT FILING DATE: 2002-04-18
;  PRIOR APPLICATION NUMBER: 60/295,031
;  PRIOR FILING DATE: 2002-05-31
;  PRIOR APPLICATION NUMBER: 60/284,773
;  PRIOR FILING DATE: 2001-04-18
;  PRIOR APPLICATION NUMBER: 60/284,764
;  PRIOR FILING DATE: 2001-04-18
;  NUMBER OF SEQ ID NOS: 88
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41
;  LENGTH: 208
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-10-125-159-41

Alignment Scores:
Pred. No.:              8.98e-19        Length:          208
Score:                  230.00          Matches:         47
Percent Similarity:     72.06%          Conservative:    2
Best Local Similarity:  69.12%          Mismatches:      18
Query Match:            30.18%          Indels:          1
DB:                     15              Gaps:            0

US-10-613-105-2 (1-132) x US-10-125-159-41 (1-208)

Qy   19 AlaProTrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLys 38
           ||||||||||||||||||            |||
Db    5 GCACCTAGAACTGGGGAAGGCTTGGGAGGGCTTGGAAAGTTCTTCAAAGCTGGAGTTTGTCTTCTACTCAAG 64

Qy   39 LysSerAlaGlnCysLeuArgTyrLys-LysProGlnSerAspTrpGlnCysPr 58
           |||||  :::   |||||||:||||||||||
Db   65 AAATGTGCCCAGTGCCTTACATACAGGGAAACCTGAGTGCCGGAGCGACTGGCAGTGGTC 124

Qy   58 OGlyLysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspTh 78
           |||||:|| |||||||||||||
Db  125 AGGGAAGGAAGAGGATGCGCCCTGACACTTGTGGCATCACATGCCTGTGTGATGTGGGGACAC 184

Qy   78 rProAsnProThrArgArgLys 85
           ||||||||||||||||||||||
Db  185 CCCCTACCCAGGGAGGAGGGAAG 206

RESULT 45
US-09-884-441-377
; Sequence 377, Application US/09884441
; Patent No. US2002011915BA1
; GENERAL INFORMATION:
;  APPLICANT: Algate, Paul A.
;  APPLICANT: Carter, Darrick
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C7
;  CURRENT APPLICATION NUMBER: US/09/884,441
;  CURRENT FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 489
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 377
;  LENGTH: 292
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-09-884-441-377

Alignment Scores:
Pred. No.:              7.09e-18        Length:          292
Score:                  224.00          Matches:         38
Percent Similarity:     100.00%         Conservative:    0
Best Local Similarity:  100.00%         Mismatches:      0
Query Match:            29.40%          Indels:          0
DB:                                     Gaps:            0

US-10-613-105-2 (1-132) x US-09-884-441-377 (1-292)
```

```
Qy    95 GlnCysLeuMetLeuAsnProProAsnPheCysGluMetAspGlyGlyGlnCysLysArgAsp 114
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 CAATGTTTGATGCTTAACCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGAC     60

Qy   115 LeuLysCysMetGlyMetCysGlyLysSerCysValSerProValLysAla 132
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TTGAAGTGTATGGGCATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 114
```

Search completed: October 24, 2004, 06:01:17
Job time : 452 secs

This Page Blank (usec)

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 24, 2004, 03:59:24 ; Search time 86 Seconds
(without alignments)
1090.978 Million cell updates/sec

Title: US-10-613-105-2
Perfect score: 762
Sequence: 1 MKSSGLFPFLVLLALGTLAP.........RDLKCCMGMCGKSCVSPVKA 132

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
DelOp   6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US10613105/runat_22102004_170441_18710/app_query.fasta_1.327
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=100 -DOCALIGN=pto -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=45
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10613105 @CGN_1_1_69 @runat_22102004_170441_18710 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	762	100.0	594	4	US-09-016-434-1271	Sequence 1271, Ap
2	458	60.1	325	4	US-09-016-434-1024	Sequence 1024, Ap
3	348	45.7	194	2	US-07-963-538B-5	Sequence 5, Appli
4	339	44.5	180	3	US-08-483-503A-2	Sequence 2, Appli
5	224	29.4	292	4	US-09-404-879A-377	Sequence 377, App
6	224	29.4	292	4	US-09-667-857-377	Sequence 377, App
7	165.5	21.7	515	4	US-09-513-999C-3787	Sequence 3787, Ap
8	165.5	21.7	643	4	US-10-000-489-41	Sequence 41, Appl
9	156	20.5	321	1	US-08-322-742-11	Sequence 11, Appl
10	152.5	20.0	478	4	US-09-023-655-1233	Sequence 1233, Ap
11	149.5	19.6	453	4	US-09-621-976-1373	Sequence 1373, Ap
12	149.5	19.6	478	4	US-09-244-111-11	Sequence 11, Appl
13	149	19.6	571	1	US-08-322-742-14	Sequence 14, Appl
14	149	19.6	2309	4	US-09-016-434-1249	Sequence 1249, Ap
15	146	19.2	93	1	US-08-304-051-15	Sequence 15, Appl
16	146	19.2	93	1	US-08-304-051-16	Sequence 16, Appl
17	146	19.2	93	5	PCT-US95-11445-15	Sequence 15, Appli
18	146	19.2	93	5	PCT-US95-11445-16	Sequence 16, Appli
19	140	18.4	177	1	US-08-379-437-3	Sequence 3, Appli
20	140	18.4	234	4	US-09-016-434-673	Sequence 673, App
21	139	18.2	177	1	US-08-379-437-5	Sequence 5, Appli
22	139	18.2	177	1	US-08-379-437-7	Sequence 7, Appli
23	134	17.6	68	3	US-07-963-538B-34	Sequence 34, Appl
24	133.5	17.5	724	3	US-09-724-864-23	Sequence 23, Appl
25	129.5	17.0	449	4	US-09-621-976-1919	Sequence 1919, A
26	129	16.9	67	2	US-07-963-538B-33	Sequence 33, Appl
27	129	16.9	80	1	US-08-304-051-7	Sequence 7, Appli
28	129	16.9	80	1	US-08-304-051-11	Sequence 11, Appl
29	129	16.9	80	5	PCT-US95-11445-7	Sequence 7, Appli
30	129	16.9	80	5	PCT-US95-11445-11	Sequence 11, Appl
31	125	16.4	66	2	US-07-963-538B-32	Sequence 32, Appl
32	124	16.3	63	2	US-07-963-538B-31	Sequence 31, Appl
33	124	16.3	89	1	US-08-304-051-12	Sequence 12, Appl
34	124	16.3	89	1	US-08-304-051-13	Sequence 13, Appl
35	124	16.3	89	5	PCT-US95-11445-12	Sequence 12, Appl
36	124	16.3	89	5	PCT-US95-11445-13	Sequence 13, Appl
37	117.5	15.4	753	4	US-09-152-060-34	Sequence 34, Appl
38	112.5	14.8	472	4	US-09-270-767-1693	Sequence 1693, Ap
39	112.5	14.8	472	4	US-09-270-767-16975	Sequence 16975, A
40	112.5	14.8	783	4	US-09-152-060-50	Sequence 50, Appl
41	109.5	14.4	437	3	US-09-383-586-29	Sequence 29, Appl
42	109.5	14.4	437	4	US-09-823-038A-29	Sequence 29, Appl
43	109	14.3	4905	1	US-07-978-895-3	Sequence 3, Appli
44	109	14.3	4905	2	US-08-475-352-3	Sequence 3, Appli
45	109	14.3	4905	2	US-09-170-699-3	Sequence 3, Appli
46	109	14.3	4975	4	US-09-610-706-3	Sequence 3, Appli
47	109	14.3	5687	3	US-09-039-268	Sequence 268, Appli
48	107.5	14.1	4157	2	US-08-162-146-2	Sequence 2, Appli
49	107.5	14.1	4157	3	US-09-314-127-2	Sequence 2, Appli
50	106	13.9	854	4	US-09-312-281C-365	Sequence 365, App
51	104	13.6	1366	2	US-08-761-248B-3	Sequence 3, Appli
52	101.5	13.3	2365	3	US-09-249-697A-5	Sequence 5, Appli
53	101.5	13.3	2365	3	US-09-363-316B-5	Sequence 5, Appli
54	100.5	13.2	3003	4	US-09-907-794A-314	Sequence 314, App
55	100.5	13.2	3003	4	US-09-905-125A-314	Sequence 314, App
56	100.5	13.2	3003	4	US-09-902-775A-314	Sequence 314, App
57	100.5	13.2	3003	4	US-09-906-700-314	Sequence 314, App
58	100.5	13.2	3003	4	US-09-903-603A-314	Sequence 314, App
59	98	12.9	1042	2	US-08-761-248B-5	Sequence 5, Appli
60	98	12.9	1071	2	US-08-761-248B-1	Sequence 1, Appli
61	97	12.7	4545	6	5183884-3	Patent No. 5183884
62	95.5	12.5	567	4	US-09-449-039A-51	Sequence 51, App
63	94.5	12.4	683	4	US-09-799-451-908	Sequence 908, App
64	94.5	12.4	6677	3	US-08-939-366-27	Sequence 27, Appl
65	94.5	12.4	6677	3	US-09-467-997-6	Sequence 6, Appli
66	94.5	12.4	13842	4	US-09-105-537-30	Sequence 30, Appl
67	94.5	12.4	36778	4	US-09-105-537-5	Sequence 5, Appli
68	94.5	12.4	38506	4	US-09-330-878-19	Sequence 19, Appli
69	94.5	12.4	38506	4	US-09-141-908-1	Sequence 1, Appli
70	94.5	12.4	38506	4	US-09-657-440-19	Sequence 19, Appli
71	94	12.3	319	4	US-07-757-022B-55	Sequence 55, Appl
72	94	12.3	390	4	US-07-757-022B-85	Sequence 85, Appl
73	94	12.3	393	4	US-07-757-022B-77	Sequence 77, Appl
74	94	12.3	417	4	US-07-757-022B-33	Sequence 33, Appl
75	94	12.3	423	4	US-07-757-022B-79	Sequence 79, Appl
76	94	12.3	468	4	US-07-757-022B-105	Sequence 105, App
77	94	12.3	469	4	US-07-757-022B-117	Sequence 117, App
78	94	12.3	471	4	US-07-757-022B-101	Sequence 101, App
79	94	12.3	471	4	US-07-757-022B-113	Sequence 113, App
80	94	12.3	516	4	US-07-757-022B-87	Sequence 87, Appl
81	94	12.3	567	4	US-07-757-022B-41	Sequence 41, Appl
82	94	12.3	576	4	US-07-757-022B-31	Sequence 31, Appl
83	94	12.3	612	4	US-07-757-022B-89	Sequence 89, Appl
84	94	12.3	621	4	US-07-757-022B-91	Sequence 91, App
85	94	12.3	621	4	US-07-757-022B-115	Sequence 115, App

```
86   94  12.3   621  4  US-07-757-022B-135   Sequence 135, App
87   94  12.3   624  4  US-07-757-022B-131   Sequence 131, App
88   94  12.3   627  4  US-07-757-022B-93    Sequence 93, Appl
89   94  12.3   660  4  US-07-757-022B-95    Sequence 95, Appl
90   94  12.3   696  4  US-07-757-022B-29    Sequence 29, Appl
91   94  12.3   714  4  US-07-757-022B-71    Sequence 71, Appl
92   94  12.3   891  4  US-07-757-022B-69    Sequence 69, Appl
93   94  12.3  1206  4  US-09-312-283C-357   Sequence 357, App
94   94  12.3  1392  4  US-07-757-022B-53    Sequence 53, Appl
95   94  12.3  3420  4  US-07-757-022B-103   Sequence 103, App
96   94  12.3  3936  4  US-07-757-022B-41    Sequence 41, Appl
97   94  12.3  4065  4  US-07-757-022B-47    Sequence 47, Appl
98   94  12.3  4086  4  US-07-757-022B-39    Sequence 39, Appl
99   94  12.3  4215  4  US-07-757-022B-61    Sequence 61, Appl
100  94  12.3  5008  4  US-07-757-022B-1     Sequence 1, Appli
```

ALIGNMENTS

RESULT 1
US-09-016-434-1271
; Sequence 1271, Application US/09016434
; Patent No. 6500938
GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/016,434
;   FILING DATE: HEREWITH
;   CLASSIFICATION:
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1271:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 594 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; IMMEDIATE SOURCE:
;   LIBRARY: GENBANK
;   CLONE: g28638
US-09-016-434-1271

Alignment Scores:
Pred. No.:   4.21e-72   Length: 594
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
```

DB:                    4                         Gaps:          0

US-10-613-105-2 (1-132) x US-09-016-434-1271 (1-594)

Qy    1  MetLysSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro   20
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19  ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGGCCCTGGGAACTCTGGCACCT   78

Qy   21  TrpAlaValGluGlyGlyLysSerPheLysAlaGlyLysValCysProProLysLysSer   40
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79  TGGGCTGTGGAAGGCGTCGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT  138

Qy   41  AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys   60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  139  GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG  198

Qy   61  LysArgCysCysProCysAspThrCysGlyIleLysCysLeuAspThrProValAspThrProAsn   80
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  199  AAGAGATGTTGTCCTGCACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC  258

Qy   81  ProThrArgArgLysProGlyLysProCysProValThrTyrGlyValCysLeuMetLeuAsn  100
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259  CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC  318

Qy  101  ProProAsnPheCysGluMetAspGlyMetGlnLysGlyArgAspLeuLysCysCysMetGly  120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319  CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC  378

Qy  121  MetCysGlyLysSerCysValSerProValLysAla  132
            |||||||||||||||||||||||||||||||||||
Db  379  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  414

RESULT 2
US-09-016-434-1024
; Sequence 1024, Application US/09016434
; Patent No. 6500938
GENERAL INFORMATION:
;  APPLICANT: Janice Au-Young
;  APPLICANT: Jeffrey J. Seilhamer
;  TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
;  TITLE OF INVENTION: PATHWAY GENE EXPRESSION
;  NUMBER OF SEQUENCES: 1490
;  CORRESPONDENCE ADDRESS:
;    ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;    STREET: 3174 PORTER DRIVE
;    CITY: PALO ALTO
;    STATE: CALIFORNIA
;    COUNTRY: USA
;    ZIP: 94304
;  COMPUTER READABLE FORM:
;    MEDIUM TYPE: Floppy disk
;    COMPUTER: IBM PC compatible
;    OPERATING SYSTEM: PC-DOS/MS-DOS
;    SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;  CURRENT APPLICATION DATA:
;    APPLICATION NUMBER: US/09/016,434
;    FILING DATE: HEREWITH
;    CLASSIFICATION:
;  PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:
;    FILING DATE:
;    CLASSIFICATION:
;  ATTORNEY/AGENT INFORMATION:
;    NAME: Zeller, Karen J.
;    REGISTRATION NUMBER: 37,071
;    REFERENCE/DOCKET NUMBER: PA-0002 US
;  TELECOMMUNICATION INFORMATION:
;    TELEPHONE: (650) 855-0555
;    TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1024:
;  SEQUENCE CHARACTERISTICS:
;    LENGTH: 325 base pairs
;    TYPE: nucleic acid

```
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;     LIBRARY: CERVNOT01
;     CLONE: 939088
US-09-016-434-1024

Alignment Scores:
Pred. No.:              3.58e-40
Score:                  458.00
Percent Similarity:     89.81%
Best Local Similarity:  88.89%
Query Match:            60.10%
DB:

US-10-613-105-2 (1-132) x US-09-016-434-1024 (1-325)

Qy    1 MetLysSerSerGlyLeuPheProPheLeuValLeuAlaLeuAlaLeuGlyThrLeuAlaPro    20
          ||||||||||||||||||||||   |||  |||||||||||||||||||||||||||||||
Db    1 ATGAAGTCCAGCGGCGGNCTCTTNCCCTTNCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCT    60

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCys-ProProLysLysSe     40
          |||||||||||||||||||||||||||||||||||  :::|||| ||||||
Db   61 TGGGCTGTGGAAGGCTCCGGAAAGTCCTTCAAAGCTGGAGTCTGTCCCTCCAAGAAATC   120

Qy   40 rAlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLy     60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 TGCCCAGTGCTTAGATACAAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAA   180

Qy   60 sLysArgCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAs     80
          |||||||||||||||||||||||||||||||||||||||||||   ||||||
Db  181 GAAGAGATGTTGCCTGACACTTGTGGCATCAAATGCCTGGATCCTNTTGACAC-CCAAA   239

Qy   80 nProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAs   100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  240 CCCAACAAGGAGGAGCT-GGGAAGTNCCAGTGACTTATGCNCAATGTTTGATGTTAAN   298

Qy  100 nProProAsnPheCysGluMet   107
          ||||||||||||||||||
Db  299 CCCCCCAA--TTTTGTGAGATG   318


RESULT 3
US-07-963-538B-5
; Sequence 5, Application US/07963538B
; Patent No. 5851983
; GENERAL INFORMATION:
;   APPLICANT: SUGIYAMA, TAKASHI
;   APPLICANT: KAMIMURA, TAKASHI
;   APPLICANT: MASUDA, KENICHI
;   APPLICANT: OKADA, MASAHIRO
;   APPLICANT: OHTSUKA, EIKO
;   APPLICANT: IMAIZUMI, ATSUSHI
;   APPLICANT: WATANABE, KUNIHITO
;   APPLICANT: SUGA, TETSUYA
;   APPLICANT: MATSUMOTO, YOHICHI
;   APPLICANT: TAKEUCHI, AKIKO
;   TITLE OF INVENTION: ELASTASE INHIBITORY POLYPEPTIDE AND
;   TITLE OF INVENTION: PROCESS FOR PRODUCTION THEREOF BY RECOMBINANT GENE
;   TITLE OF INVENTION: TECHNOLOGY
;   NUMBER OF SEQUENCES: 36
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: COOLEY GODWARD LLP
;     STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
;     CITY: PALO ALTO
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 94306-2155
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
```

```
;     APPLICATION NUMBER: US/07/963,538B
;     FILING DATE: 20-OCT-1992
;     CLASSIFICATION: 530
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 07/843,359
;     FILING DATE: 25-FEB-1992
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 07/408,483
;     FILING DATE: 22-AUG-1989
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: JP 4-212399
;     FILING DATE: 17-JUL-1992
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: JP 4-212398
;     FILING DATE: 17-JUL-1992
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: JP 3-355553
;     FILING DATE: 24-DEC-1991
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: JP 62-330219
;     FILING DATE: 28-DEC-1987
;   ATTORNEY/AGENT INFORMATION:
;     NAME: NEELEY PH.D., RICHARD L.
;     REGISTRATION NUMBER: 30,092
;     REFERENCE/DOCKET NUMBER: TEJN-005/02US
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 415-843-5070
;     TELEFAX: 415-857-0663
;     TELEX: 380816 COOLEY PA
;   INFORMATION FOR SEQ ID NO:     5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 194 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: double
;       TOPOLOGY: linear
;     MOLECULE TYPE: DNA (synthetic)
US-07-963-538B-5

Alignment Scores:
Pred. No.:              8.47e-29           Length:         194
Score:                  348.00             Matches:        59
Percent Similarity:     100.00%            Conservative:   0
Best Local Similarity:  100.00%            Mismatches:     0
Query Match:            45.67%             Indels:         0
DB:                     2                  Gaps:           0

US-10-613-105-2 (1-132) x US-07-963-538B-5 (1-194)

Qy   74 AspProValAspThrProAsnProThrArgArgLysProGlyLysCysProValThrTyr    93
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GATCCGGTCGACACCCCGAACCCGACGCGTGTAAACCGGGTAAATGTCCGGTTACATAT    60

Qy   94 GlyGlnCysLeuMetLeuAsnProProAsnPheCysGluMetCysAspGlyGlnCysLysArg   113
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GGTCAGTGTCTGATGCTGAACCCGCCGAACTTCTGTGAAATGGACGGTCAGTGTAAACGA   120

Qy  114 AspLeuLysCysCysMetGlyMetCysGlyLysSerCysValSerProValLysAla   132
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GATCTGAAATGTTGTATGGGTATGTGTGGTAAATCTTGTGTTTCTCCGGTTAAAGCA   177


RESULT 4
US-08-483-503A-2
; Sequence 2, Application US/08483503A
; Patent No. 6017880
; GENERAL INFORMATION:
;   APPLICANT: Eisenberg, Stephen
;   APPLICANT: Wahl, Sharon M.
;   APPLICANT: Thompson, Robert C.
;   APPLICANT: Dripps, David J.
;   TITLE OF INVENTION: INHIBITION OF RETROVIRUS INFECTION
;   NUMBER OF SEQUENCES: 4
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
```

```
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,503A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/209,040
; FILING DATE: 09-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,369
; FILING DATE: 09-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barker, M. Paul
; REGISTRATION NUMBER: 32,013
; REFERENCE/DOCKET NUMBER: 04189.0084-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-483-503A-2

Alignment Scores:
Pred. No.:                6.9e-28       Length:          180
Score:                    339.00        Matches:          58
Percent Similarity:       96.67%        Conservative:      0
Best Local Similarity:    96.67%        Mismatches:        2
Query Match:              44.49%        Indels:            0
DB:                                     Gaps:              0

US-10-613-105-2 (1-132) x US-08-483-503A-2 (1-180)

Qy  73 LeuAspProValAspThrProAsnProThrArgArgLysProGlyLysCysProValThr  92
         |||||||||||||||          ||||||||||||||||||||||||||||||||||||
Db   1 CTGGATCCTGTTGACACCCCAACACAGGAGGAAGCCTGGGAAGTGCCCAGTGACT       60

Qy  93 TyrGlyGlnCysLeuMetLeuAsnProProAsnPheCysGluMetGlyGlnCysLys     112
         ||||||||||||||||||||||||||||||||||||
Db  61 TATGGCCAATGTTGATGCTTAACCCCCCAATTTCTGTGAGATGGGCCAGTGCAAG      120

Qy 113 ArgAspLeuLysCysMetGlyMetCysValSerProValLysAla                 132
         |||||||||||||||||||||||||||||||||||
Db 121 CGTGACTTGAAGTGTTGCATGGGCATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  180

RESULT 5
US-09-404-879A-377
; Sequence 377, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 377
; LENGTH: 292
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-404-879A-377

Alignment Scores:
Pred. No.:                2.23e-15      Length:          292
Score:                    224.00        Matches:          38
Percent Similarity:       100.00%       Conservative:      0
Best Local Similarity:    100.00%       Mismatches:        0
Query Match:              29.40%        Indels:            0
DB:                                     Gaps:              0

US-10-613-105-2 (1-132) x US-09-404-879A-377 (1-292)

Qy  95 GlnCysLeuMetLeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysArgAsp  114
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CAATGTTGATGCTTAACCCCCCAATTTCTGTGAGATGGGCCAGTGCAAGCGTGAC       60

Qy 115 LeuLysCysMetGlyMetCysGlyLysSerGlyLysSerProValLysAla           132
         |||||||||||||||||||||||||||||||||||
Db  61 TTGAAGTGTTGCATGGGCATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT        114

RESULT 6
US-09-667-857-377
; Sequence 377, Application US/09667857
; Patent No. 6699664
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C5
; CURRENT APPLICATION NUMBER: US/09/667,857
; CURRENT FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 377
; LENGTH: 292
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-667-857-377

Alignment Scores:
Pred. No.:                2.23e-15      Length:          292
Score:                    224.00        Matches:          38
Percent Similarity:       100.00%       Conservative:      0
Best Local Similarity:    100.00%       Mismatches:        0
Query Match:              29.40%        Indels:            0
DB:                                     Gaps:              0

US-10-613-105-2 (1-132) x US-09-667-857-377 (1-292)

Qy  95 GlnCysLeuMetLeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysArgAsp  114
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CAATGTTGATGCTTAACCCCCCAATTTCTGTGAGATGGGCCAGTGCAAGCGTGAC       60

Qy 115 LeuLysCysMetGlyMetCysGlyLysSerGlyLysSerProValLysAla           132
         |||||||||||||||||||||||||||||||||||
Db  61 TTGAAGTGTTGCATGGGCATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT        114

RESULT 7
US-09-513-999C-3787
; Sequence 3787, Application US/09513999C
```

; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 3787
; LENGTH: 515
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 57..428
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 57..146
; OTHER INFORMATION: score 9.9
; OTHER INFORMATION: seq LLFGFTLVSGTGA/EK
US-09-513-999C-3787

Alignment Scores:
Pred. No.:   515
Score: 8.04e-09   Matches: 44
Percent Similarity: 165.50   Conservative: 14
Best Local Similarity: 45.67%   Mismatches: 44
Query Match: 34.65%   Indels: 25
DB: 21.72%   Gaps: 7
                        4

US-10-613-105-2 (1-132) x US-09-513-999C-3787 (1-515)

Qy   11 ValLeuLeuAlaLeuGlyThrLeuAlaProTrpAlaValGluGlySerGlyLysSerPhe 30
          ::::|||||::::|||
Db   93 CTCCCTCCTCAGCCTGCTGTTCGGCTTCACCCTAGTCTCAGGCACAGGAGCAGAG--- 149

Qy   31 LysAlaGlyValCysProProLysSerAlaGln---CysLeuArgTyrLysLysPro 49
          |||| :::::: ||| :::
Db   150 AAGACTGGCGTGTGCCCGAGCTCCAGGCTGACCAGAACTGCACGCAA---------- 197

Qy   50 GluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysProAspThrCysGly 69
          ||||| ::|||| :: ||||||
Db   198 GAGTGCGTCTCGGACAGCGAATGCGCCGACAACCTCAAGTGCTGCGAGCGCGGCTGTGCC 257

Qy   70 IleLysCysLeuAspProValAspThrProAsnProThrArgArgLysProGlyLysCys 89
          ||||||||
Db   258 ACCTTCTGCTCTCTGCCCAATGAT-----------AAGGAGGGTTCCTGC 296

Qy   90 ProValThrTyrGlyGlnCysLeuMetLeuAsnProProAsnPhe------------ 104
          ::: ::|||| |||
Db   297 CCCCAG------------GTGAACATTAACTTTCCCCAGCTCGGCCTCTGTCGGGAC 341

Qy   105 ---CysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysMetGlyMetCysGly 123
          :::||||||||||:::|||||
Db   342 CAGTGCCAGGTGGACAGCCAGTGTCCTGGCCAGATGAAATGCTGCCGCAATGGCTGTGGG 401

Qy   124 Lys--SerCysValSerPro 129
          ||||||||||:::|||
Db   402 AAGGTGTCCTGTGTCACTCCC 422

RESULT 8
US-10-000-489-41
; Patent No. 6794363
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; Sequence 41. Application US/10000489

; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF .
; FILE REFERENCE: 91.US6.DIV
; CURRENT APPLICATION NUMBER: US/10/000,489
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 41
; LENGTH: 643
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..90
; NAME/KEY: CDS
; LOCATION: 91..462
; NAME/KEY: 3'UTR
; LOCATION: 463..643
; NAME/KEY: polyA_signal
; LOCATION: 607..612
; NAME/KEY: polyA_site
; LOCATION: 628..643
US-10-000-489-41

Alignment Scores:
Pred. No.:   643
Score: 1.09e-08   Matches: 44
Percent Similarity: 165.50   Conservative: 14
Best Local Similarity: 45.67%   Mismatches: 44
Query Match: 34.65%   Indels: 25
DB: 21.72%   Gaps: 7
                        4

US-10-613-105-2 (1-132) x US-10-000-489-41 (1-643)

Qy   11 ValLeuLeuAlaLeuGlyThrLeuAlaProTrpAlaValGluGlySerGlyLysSerPhe 30
          ::::|||||::::|||
Db   127 CTCCCTCCTCAGCCTGCTGTTCGGCTTCACCCTAGTCTCAGGCACAGGAGCAGAG--- 183

Qy   31 LysAlaGlyValCysProProLysSerAlaGln---CysLeuArgTyrLysLysPro 49
          |||| :::::: ||| :::
Db   184 AAGACTGGCGTGTGCCCGAGCTCCAGGCTGACCAGAACTGCACGCAA---------- 231

Qy   50 GluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysProAspThrCysGly 69
          ||||| ::|||| :: ||||||
Db   232 GAGTGCGTCTCGGACAGCGAATGCGCCGACAACCTCAAGTGCTGCGAGCGCGGCTGTGCC 291

Qy   70 IleLysCysLeuAspProValAspThrProAsnProThrArgArgLysProGlyLysCys 89
          ||||||||
Db   292 ACCTTCTGCTCTCTGCCCAATGAT-----------AAGGAGGGTTCCTGC 330

Qy   90 ProValThrTyrGlyGlnCysLeuMetLeuAsnProProAsnPhe------------ 104
          ::: ::|||| |||
Db   331 CCCCAG------------GTGAACATTAACTTTCCCCAGCTCGGCCTCTGTCGGGAC 375

Qy   105 ---CysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysMetGlyMetCysGly 123
          :::||||||||||:::|||||
Db   376 CAGTGCCAGGTGGACAGCCAGTGTCCTGGCCAGATGAAATGCTGCCGCAATGGCTGTGGG 435

Qy   124 Lys--SerCysValSerPro 129
          ||||||||||:::|||
Db   436 AAGGTGTCCTGTGTCACTCCC 456

RESULT 9
US-08-322-742-11
; Sequence 11, Application US/08322742
; Patent No. 5688641
; GENERAL INFORMATION:
;   APPLICANT: Sager, Ruth
;   TITLE OF INVENTION: CANCER DIAGNOSIS AND THERAPY
;   NUMBER OF SEQUENCES: 19
;   CORRESPONDENCE ADDRESS:
;   ADDRESSE: Fish & Richardson
;   STREET: 225 Franklin Street
;   CITY: Boston
;   STATE: Massachusetts
;   COUNTRY: U.S.A.
;   ZIP: 02110-2804
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;   COMPUTER: IBM PS/2 Model 50Z or 55SX
;   OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
;   SOFTWARE: WordPerfect (Version 5.1)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/322,742
;   FILING DATE:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/938,823
;   FILING DATE: September 1, 1992
;   APPLICATION NUMBER: 07/844,296
;   FILING DATE: February 28, 1992
;   APPLICATION NUMBER: 07/552,216
;   FILING DATE: February 28, 1991
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Fraser, Janis K.
;   REGISTRATION NUMBER: 34,819
;   REFERENCE/DOCKET NUMBER: 00530/048003
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (617) 542-5070
;   TELEFAX: (617) 542-8906
;   TELEX: 200154
; INFORMATION FOR SEQ ID NO: 11:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 321
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
US-08-322-742-11

Alignment Scores:
Pred. No.:   4.25e-08   Length: 321
Score: 156.00   Matches: 33
Percent Similarity: 39.78%   Conservative: 4
Best Local Similarity: 35.48%   Mismatches: 36
Query Match: 20.47%   Indels: 20
DB: 1   Gaps: 2

US-10-613-105-2 (1-132) x US-08-322-742-11 (1-321)

Qy   57 CysProGly---------LysLysArgCysCysProAspThrCysGlyIleLysCys 72
                         ::: ||| ||| ||| |||
Db   33 TGTCACGGGAGTTCCTGTTAAAGGTCAAGACCCCTGTCAAAGGCCGTGTTCCATCAATGGA 92

Qy   73 LeuAspProVal--- 76
            ||||||||||
Db   93 CAAGATCCCGTTAAAGGACAACAGTTCAGTTAAAGGTCAAGATAAAGTCAAGACGCAAGAG 152

Qy   77 AspThrProAsnProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCys 96
                        ||| ::: |||||
Db   153 CCAGTCAAAGGTCCAGTCTCCACTAAGCCTGCTCCTGCCCATTATCTTGATCCCGGTGC 212

Qy   97 LeuMetLeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLys 116
            |||||||| ||| ||| ||||
Db   213 GCCATGTTGAATCCTCCTAACCGCTGCTTGAAAGATACTGACTGCCCAGGAATCAAGAAG 272

Qy   117 CysCysMetGlyMetCysGlyLysLysSerCysValSerPro 129

Db   273 TGCTGTGAAGGCTCTTGCGGGATGGCCTGTTCGTTCCC 311
          ||||||| ||| |||||| |||||| :::|||

RESULT 10
US-09-023-655-1233
; Sequence 1233, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
;   APPLICANT: Cocks, Benjamin G.
;   APPLICANT: Suban G. Stuart
;   APPLICANT: Jeffrey J. Seilhamer
;   TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
;   TITLE OF INVENTION: EXPRSSION
;   NUMBER OF SEQUENCES: 1508
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/023,655
;   FILING DATE: HEREWITH
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0001 US
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1233:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 478 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;   LIBRARY: GENBANK
;   CLONE: g28711
US-09-023-655-1233

Alignment Scores:
Pred. No.:   1.74e-07   Length: 478
Score: 152.50   Matches: 47
Percent Similarity: 41.48%   Conservative: 9
Best Local Similarity: 34.81%   Mismatches: 55
Query Match: 20.01%   Indels: 25
DB: 4   Gaps: 5

US-10-613-105-2 (1-132) x US-09-023-655-1233 (1-478)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAla-- 19
            :::::|::| ||| ||
Db   3 ATGAGGGCCAGCAGCTTCTTGATCGTGTGGTGTTCCTCATCGCTGGGACGCTGGTGTCTA 62

Qy   20 ProTrpAlaValGluGlyLysSerGlyLysSerPheLysAlaGlyValCysProProLysLys 39
            ||||| ||| ||
Db   63 GAGGCAGCTGTCACGGGA--------------------------GTTCCTGTTAAAGG- 94

Qy   40 SerAlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrp-GlnCysProGl 59
            |||||| |||
Db   95 TCAAGACACTGTCAAAGG------------------CCGTGTTCCATTCAATGGACAAGATCCCGT 142

```
Db  271 CCCCAG------------GTGAACATTAACTTTCCCCAGCTCGGCCTCTGTCGGGAC 315

Qy  105 ---CysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMetCysGly 123
            :::|||||||       ::|||||||

Db  316 CAGTGCCAGTGGACAGCCAGTGCCTGGCCAGATGAAATGCTGCCCAATGCGCTGTGGG 375

Qy  124 Lys---SerCysValSerPro 129

Db  376 AAGGTGTCCTGTGTCACTCCC 396
```

RESULT 12
US-09-244-111-11
; Sequence 11, Application US/09244111
; Patent No. 6566498
; GENERAL INFORMATION:
; APPLICANT: Ni, et al.
; TITLE OF INVENTION: Human Serine Protease and Serpin Polypeptides
; FILE REFERENCE: PF391
; CURRENT APPLICATION NUMBER: US/09/244,111
; CURRENT FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: 60/073,961
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(249)
US-09-244-111-11

Alignment Scores:
Pred. No.:	3.63e-07	Length:	478
Score:	149.50	Matches:	38
Percent Similarity:	54.95%	Conservative:	12
Best Local Similarity:	41.76%	Mismatches:	34
Query Match:	19.62%	Indels:	7
DB:	4	Gaps:	5

US-10-613-105-2 (1-132) x US-09-244-111-11 (1-478)

```
Qy    1 MetLysSerGlyLeuProPheProLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20

Db   19 ATGGGGTCCAGCAGCTTCTTGGTCTCCATGGTGTCTCTCGTTCTTGTGACCCTGGTG--- 75

Qy   21 TrpAlaValGluGlySerGlyLysSerPhe---LysAlaGlyValCysProLysGlys 39

Db   76 ---GCTGTGGAAGGAGTTAAAGAGGGTATAGAGAAAGCAGGGGGTTGC---CCAGCTGAC 129

Qy   40 SerAlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGly 59

Db  130 AACGTACGTCGTTCAGTCCCATCCTCCCCCAGTGTCACACAGGACTGTCTGGGG 189

Qy   60 LysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrPro 79

Db  190 GAAAGGAAGTGTTGTTACCTGCACTGTGGCTTCCAAGTGTGTGATTCCTGTG------AAG 243

Qy   80 AsnProThrArgArgLysProGly---LysCys 89

Db  244 AACTGAGAAGGAGGAAACAAGGATGAAGATGT 276
```

RESULT 13
US-08-322-742-14
; Sequence 14, Application US/08322742
; Patent No. 5688641
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; TITLE OF INVENTION: CANCER DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson

```
Qy   59 yLysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrPr 79
            :||   ||        |||

Db  143 TAAAGGACAA------------GTTTCAGTTAAAGTCAAGATAAAGTCAAAGCGCA 187

Qy   79 oAsnProThrArg---------ArgLysProGlyLysCysProValThrTyrGlyGl 95
            |||  ::

Db  188 AGAGCCAGTCAAAGGTCCAGTCTCCACTAAGCCTGGCCTGGCCCATTATCTTGATCCG 247

Qy   95 nCysLeuMetLeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLe 115
            |||||||||||

Db  248 GTGCGCCATGTTGAATCCCCTAACCGCGTCTTGAAAGATACTGACTGCCCAGGAATCAA 307

Qy  115 uLysCysCysMetGlyMetCysGlyLysSerCysValSerPro 129
            :::|||    ||

Db  308 GAAGTGCTGTGAAGGCTCTTGCGGATGGCCTGTTTCGTTCCC 350
```

RESULT 11
US-09-621-976-1373
; Sequence 1373, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1373
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 32..277
; NAME/KEY: sig_peptide
; LOCATION: 32..121
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 9.89999961853027
; OTHER INFORMATION: seq LLFGFTLVSGTGA/EK
US-09-621-976-1373

Alignment Scores:
Pred. No.:	3.37e-07	Length:	453
Score:	149.50	Matches:	44
Percent Similarity:	45.67%	Conservative:	14
Best Local Similarity:	34.65%	Mismatches:	44
Query Match:	19.62%	Indels:	26
DB:	4	Gaps:	7

US-10-613-105-2 (1-132) x US-09-621-976-1373 (1-453)

```
Qy   11 ValLeuLeuAlaLeuGlyThrLeuAlaProTrpAlaValGluGlySerGlyLysSerPhe 30
            :||:::||:::||

Db   68 CTCCTCCTCAGCCTGCGTCGTTCGGCTTCACCCTAGTCTCAGGCACAGGAGCAGAG--- 124

Qy   31 LysAlaGlyValCysProLysProLysSerProAlaGln---CysLeuArgTyrLysLysPro 49
            ||||||||||||:::

Db  125 AAGACTGGGCGTGTGCCCGAGCTCCAGGCTGACCAGAACTGCACCCAA----------- 172

Qy   50 GluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysProAspThrCysGly 69
            ||||||     :||||

Db  173 GAGTGCGTCTCGGACAGCAGGAATGCGCCGACAACCTCAAGTCTCCAGGCGGGTCTGCA 232

Qy   70 IleLysCysLeuAspProValAspThrProAsnProThrArgArgLysProGlyLysCys 89
            ||||||     :::

Db  233 CCTTCTGCTCTCTG-CCCAATGAT----------------AAGGAGGGTTCCTGC 270

Qy   90 ProValThrTyrGlyGlnCysLeuMetLeuAsnProProAsnPhe------------ 104
            ::: ||     |||
```

```
;           STREET:  225 Franklin Street
;           CITY:  Boston
;           STATE:  Massachusetts
;           COUNTRY:  U.S.A.
;           ZIP:  02110-2804
;     COMPUTER READABLE FORM:
;           MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;           COMPUTER:  IBM PS/2 Model 50Z or 55SX
;           OPERATING SYSTEM:  IBM P.C. DOS (Version 3.30)
;           SOFTWARE:  WordPerfect (Version 5.1)
;     CURRENT APPLICATION DATA:
;           APPLICATION NUMBER:  US/08/322,742
;           FILING DATE:
;     PRIOR APPLICATION DATA:
;           APPLICATION NUMBER:  07/938,823
;           FILING DATE:  September 1, 1992
;           APPLICATION NUMBER:  07/844,296
;           FILING DATE:  February 28, 1992
;           APPLICATION NUMBER:  07/552,216
;           FILING DATE:  February 28, 1991
;     ATTORNEY/AGENT INFORMATION:
;           NAME:  Fraser, Janis K.
;           REGISTRATION NUMBER:  34,819
;           REFERENCE/DOCKET NUMBER:  00530/048003
;     TELECOMMUNICATION INFORMATION:
;           TELEPHONE:  (617) 542-5070
;           TELEFAX:  (617) 542-8906
;           TELEX:  200154
; INFORMATION FOR SEQ ID NO: 14:
;     SEQUENCE CHARACTERISTICS:
;           LENGTH:  571
;           TYPE:  nucleic acid
;           STRANDEDNESS:  double
;           TOPOLOGY:  linear

US-08-322-742-14

Alignment Scores:
Pred. No.:                5.25e-07        Length:          571
Score:                    149.00          Matches:         28
Percent Similarity:       55.00%          Conservative:    5
Best Local Similarity:    46.67%          Mismatches:      25
Query Match:              19.55%          Indels:          2
DB:                                       Gaps:            1

US-10-613-105-2 (1-132) x US-08-322-742-14 (1-571)

Qy   70  IleLysCysLeuAspProValAspThrProAsnProThrArgArgLysProGlyLysCys  89
             ||:|:||       ||||||||||||||||||||||
Db  193  GTCAAGGCCACAAGAGACCAGTCAAAGTCCAGTCCACT-----AAGCCTGGCTCCTGC  246

Qy   90  ProValThrTyrGlyGlnCysLeuMetLeuAsnProProAsnPheCysGluMetAspGly  109
             ||:|:|          ::|||:|   ||||||||||||||||||||
Db  247  CCCATTATCTTGATCCGGTGCGCCATGTTGAATCCTCCTAACCGTGCTTGAAAGATACT  306

Qy  110  GlnCysLysArgAspLeuLysCysMetGlyMetCysGlyLysSerCysValSerPro  129
             ||||||         |||:||    ::|||||
Db  307  GACTGCCCAGGAATCAAGAAGTGCTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCC  366
```

RESULT 14
US-09-016-434-1249
; Sequence 1249, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
;   APPLICANT: Janice Au-Young
;   APPLICANT: Jeffrey J. Seilhamer
;   TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
;   TITLE OF INVENTION: PATHWAY GENE EXPRESSION
;   NUMBER OF SEQUENCES: 1490
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA

```
;           COUNTRY:  USA
;           ZIP:  94304
;     COMPUTER READABLE FORM:
;           MEDIUM TYPE:  Floppy disk
;           COMPUTER:  IBM PC compatible
;           OPERATING SYSTEM:  PC-DOS/MS-DOS
;           SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;     CURRENT APPLICATION DATA:
;           APPLICATION NUMBER:  US/09/016,434
;           FILING DATE:  HEREWITH
;           CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;           APPLICATION NUMBER:
;           FILING DATE:
;           CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;           NAME:  Zeller, Karen J.
;           REGISTRATION NUMBER:  37,071
;           REFERENCE/DOCKET NUMBER:  PA-0002 US
;     TELECOMMUNICATION INFORMATION:
;           TELEPHONE:  (650) 855-0555
;           TELEFAX:  (650) 845-4166
; INFORMATION FOR SEQ ID NO:  1249:
;     SEQUENCE CHARACTERISTICS:
;           LENGTH:  2309 base pairs
;           TYPE:  nucleic acid
;           STRANDEDNESS:  single
;           TOPOLOGY:  linear
;     IMMEDIATE SOURCE:
;           LIBRARY:  GENBANK
;           CLONE:  g190337

US-09-016-434-1249

Alignment Scores:
Pred. No.:                3.67e-06        Length:          2309
Score:                    149.00          Matches:         38
Percent Similarity:       43.97%          Conservative:    13
Best Local Similarity:    32.76%          Mismatches:      52
Query Match:              19.55%          Indels:          13
DB:                                       Gaps:            4

US-10-613-105-2 (1-132) x US-09-016-434-1249 (1-2309)

Qy   21  TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer  40
             |||       |||       ::|||
Db 1394  TGGCAGTGAAGCAGAGAGCTTACTGGGTATAAAATGTGGCTCGTTCTCTTTTAACAGT  1453

Qy   41  AlaGlnCys-----LeuArgTyrLysLysProGluCysGlnSerAspTrp-GlnCysPr  58
             :::  |||        ||||:|                |||
Db 1454  TCC---TGTTAAAGGTCAAGACACTGTCAAAGGCCGTTCCATTCAATGGACAAGATCC  1510

Qy   58  oGlyLysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspTh  78
             :::|||            ||:|
Db 1511  CGTTAAAGGACAA------GTTTCAGTAAAGGTCAAGATAAAGTCAAGAGTCAAAGC  1555

Qy   78  rProAsnProThrArg-------------ArgLysProGlyLysCysProValThrTyrGl  94
             :::                                         |||:::
Db 1556  GCAAGAGCAGTCAAAGTCCAGTCCACTAAGCCTGGCTCCTGCCCATTATCTTGAT  1615

Qy   94  yGlnCysLeuMetLeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAs  114
             :::||||||||||||||||       ||||||
Db 1616  CCGGTGCGCCATGTTGAATCCCCTAACCGTGCTTGAAAGATACTGACTGCCCAGGAAT  1675

Qy  114  pLeuLysCysCysMetGlyMetCysGlyLysSerCysValSerPro  129
             |||      :::|||
Db 1676  CAAGAAGTGCTGTGAAGGCTCTTGCGGGATGGCCTGTTTCGTTCCC  1721
```

RESULT 15
US-08-304-051-15
; Sequence 15, Application US/08304051
; Patent No. 5633227
; GENERAL INFORMATION:
;   APPLICANT: Muller, Daniel K.; Brownell, Elise; Delaria, Katherine A.

```
    TITLE OF INVENTION: Secretory Leukocyte Protease Inhibitor as an
    TITLE OF INVENTION: Inhibitor of Tryptase
    NUMBER OF SEQUENCES: 21
    CORRESPONDENCE ADDRESS:
       ADDRESSEE: Miles Inc.
       STREET: 400 Morgan Lane
       CITY: West Haven
       STATE: Connecticut
       COUNTRY: USA
       ZIP: 06516
    COMPUTER READABLE FORM:
       MEDIUM TYPE: Diskette, 3.50 inch, 1,300 Kb storage
       COMPUTER: Apple Macintosh
       OPERATING SYSTEM: System 7.1
       SOFTWARE: Word Perfect 3.0a
    CURRENT APPLICATION DATA:
       APPLICATION NUMBER: US/08/304,051
       FILING DATE: 12-Sept-1994
       CLASSIFICATION: 514
    PRIOR APPLICATION DATA: No. 5633227e
    ATTORNEY/AGENT INFORMATION:
       NAME: William F. Gray
       REGISTRATION NUMBER: 31018
       REFERENCE/DOCKET NUMBER: MWH 322
    TELECOMMUNICATION INFORMATION:
       TELEPHONE: (203) 937-2712
       TELEFAX: (203) 937-2795
       TELEX: 221949 MILES UR
    INFORMATION FOR SEQ ID NO: 15:
    SEQUENCE CHARACTERISTICS:
       LENGTH: 93 bases
       TYPE: nucleic acid
       STRANDEDNESS: single
       TOPOLOGY: linear
    MOLECULE TYPE:
    MOLECULE TYPE: - Other nucleic Acid: synthetic
    HYPOTHETICAL: No
    ANTI-SENSE: No
    FEATURE:
       NAME/KEY: Primer for Polymerase Chain Reaction used to make mutein
       NAME/KEY: of human SLPI.
    US-08-304-051-15

Alignment Scores:
Pred. No.:                               Length:        93
Score:                 8.76e-08          Matches:       26
Percent Similarity:    146.00            Conservative:  0
Best Local Similarity: 96.30%            Mismatches:    1
Query Match:           96.30%            Indels:        0
DB:                    19.16%            Gaps:          0
                       1

US-10-613-105-2 (1-132) x US-08-304-051-15 (1-93)

Qy  25 GlySerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeu 44
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12 GGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCTGCCCAGTGCAAGA 71

Qy  45 ArgTyrLysLysProGluCys 51
        |||||||||||||||||
Db  72 AGATACAAGAAACCTGAGTGC 92

RESULT 16
US-08-304-051-16
; Sequence 16, Application US/08304051
; Patent No. 5633227
    GENERAL INFORMATION:
       APPLICANT: Muller, Daniel K.; Brownell, Elise; Delaria, Katherine A.
       TITLE OF INVENTION: Secretory Leukocyte Protease
       TITLE OF INVENTION: Inhibitor of Tryptase as an
       NUMBER OF SEQUENCES: 21
    CORRESPONDENCE ADDRESS:
       ADDRESSEE: Miles Inc.
       STREET: 400 Morgan Lane
```

```
       CITY: West Haven
       STATE: Connecticut
       COUNTRY: USA
       ZIP: 06516
    COMPUTER READABLE FORM:
       MEDIUM TYPE: Diskette, 3.50 inch, 1,300 Kb storage
       COMPUTER: Apple Macintosh
       OPERATING SYSTEM: System 7.1
       SOFTWARE: Word Perfect 3.0a
    CURRENT APPLICATION DATA:
       APPLICATION NUMBER: US/08/304,051
       FILING DATE: 12-Sept-1994
       CLASSIFICATION: 514
    PRIOR APPLICATION DATA: No. 5633227e
    ATTORNEY/AGENT INFORMATION:
       NAME: William F. Gray
       REGISTRATION NUMBER: 31018
       REFERENCE/DOCKET NUMBER: MWH 322
    TELECOMMUNICATION INFORMATION:
       TELEPHONE: (203) 937-2712
       TELEFAX: (203) 937-2795
       TELEX: 221949 MILES UR
    INFORMATION FOR SEQ ID NO: 16:
    SEQUENCE CHARACTERISTICS:
       LENGTH: 93 bases
       TYPE: nucleic acid
       STRANDEDNESS: single
       TOPOLOGY: linear
    MOLECULE TYPE:
    MOLECULE TYPE: - Other nucleic Acid: synthetic
    HYPOTHETICAL: No
    ANTI-SENSE: No
    FEATURE:
       NAME/KEY: Primer for Polymerase Chain Reaction used to make mutein
       NAME/KEY: of human SLPI.
    US-08-304-051-16

Alignment Scores:
Pred. No.:                               Length:        93
Score:                 8.76e-08          Matches:       26
Percent Similarity:    146.00            Conservative:  0
Best Local Similarity: 96.30%            Mismatches:    1
Query Match:           96.30%            Indels:        0
DB:                    19.16%            Gaps:          0
                       1

US-10-613-105-2 (1-132) x US-08-304-051-16 (1-93)

Qy  25 GlySerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeu 44
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12 GGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCTGCCCAGTGCAAG 71

Qy  45 ArgTyrLysLysProGluCys 51
        |||||||||||||||||
Db  72 AGATACAAGAAACCTGAGTGC 92

RESULT 17
PCT-US95-11445-15
; Sequence 15, Application PC/TUS9511445
    GENERAL INFORMATION:
       APPLICANT: Muller, Daniel K.; Brownell, Elise; Delaria,
       APPLICANT: Katherine A.
       TITLE OF INVENTION: Secretory Leukocyte Protease
       TITLE OF INVENTION: Inhibitor as an Inhibitor of Tryptase
       NUMBER OF SEQUENCES: 21
    CORRESPONDENCE ADDRESS:
       ADDRESSEE: Bayer Corporation
       STREET: 400 Morgan Lane
       CITY: West Haven
       STATE: Connecticut
       COUNTRY: USA
       ZIP: 06516
    COMPUTER READABLE FORM:
       MEDIUM TYPE: Diskette, 3.50 inch, 1,300 Kb storage
```

```
            COMPUTER: Apple Macintosh
    OPERATING SYSTEM: System 7.1
            SOFTWARE: Word Perfect 3.0a
CURRENT APPLICATION DATA:
  APPLICATION NUMBER:  PCT/US95/11445
         FILING DATE:  11 September 1995
      CLASSIFICATION:
PRIOR APPLICATION DATA:
  APPLICATION NUMBER:  08/304,051
         FILING DATE:  12 SEPTEMBER 1994
ATTORNEY/AGENT INFORMATION:
                NAME:  William F. Gray
 REGISTRATION NUMBER:  31018
REFERENCE/DOCKET NUMBER:  MWH 322P1
TELECOMMUNICATION INFORMATION:
           TELEPHONE:  (203) 937-2712
             TELEFAX:  (203) 937-5492
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
              LENGTH:  93 bases
                TYPE:  nucleic acid
         STRANDEDNESS:  single
            TOPOLOGY:  linear
MOLECULE TYPE:
       MOLECULE TYPE:  - Other nucleic Acid: synthetic
        HYPOTHETICAL:  No
          ANTI-SENSE:  No
FEATURE:
            NAME/KEY:  Primer for Polymerase Chain Reaction used to
            NAME/KEY:  make mutein of human SLPI.
PCT-US95-11445-15

Alignment Scores:
           Pred. No.:  8.76e-08         Length:      93
               Score:  146.00           Matches:     26
 Percent Similarity:   96.30%           Conservative:  0
Best Local Similarity:  96.30%           Mismatches:    1
         Query Match:  19.16%           Indels:        0
                  DB:  5                Gaps:          0

US-10-613-105-2 (1-132) x PCT-US95-11445-15  (1-93)

Qy     25 GlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSerAlaGlnCysLeu 44
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db     12 GGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCTGCCCAGTGCAGA 71

Qy     45 ArgTyrLysLysProGluCys 51
          |||||||||||||||
Db     72 AGATACAAGAAACCTGAGTGC 92

RESULT 18
PCT-US95-11445-16
; Sequence 16, Application PC/TUS9511445
; GENERAL INFORMATION:
           APPLICANT:  Muller, Daniel K.; Brownell, Elise; Delaria,
           APPLICANT:  Katherine A.
   TITLE OF INVENTION:  Secretory Leukocyte Protease
   TITLE OF INVENTION:  Inhibitor as an Inhibitor of Tryptase
NUMBER OF SEQUENCES:  21
CORRESPONDENCE ADDRESS:
          ADDRESSEE:  Bayer Corporation
             STREET:  400 Morgan Lane
               CITY:  West Haven
              STATE:  Connecticut
            COUNTRY:  USA
                ZIP:  06516
COMPUTER READABLE FORM:
         MEDIUM TYPE:  Diskette, 3.50 inch, 1,300 Kb storage
            COMPUTER:  Apple Macintosh
    OPERATING SYSTEM:  System 7.1
            SOFTWARE:  Word Perfect 3.0a
CURRENT APPLICATION DATA:
  APPLICATION NUMBER:  PCT/US95/11445
```

```
         FILING DATE:  11 September 1995
      CLASSIFICATION:
PRIOR APPLICATION DATA:
  APPLICATION NUMBER:  08/304,051
         FILING DATE:  12 SEPTEMBER 1994
ATTORNEY/AGENT INFORMATION:
                NAME:  William F. Gray
 REGISTRATION NUMBER:  31018
REFERENCE/DOCKET NUMBER:  MWH 322P1
TELECOMMUNICATION INFORMATION:
           TELEPHONE:  (203) 937-2712
             TELEFAX:  (203) 937-5492
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
              LENGTH:  93 bases
                TYPE:  nucleic acid
         STRANDEDNESS:  single
            TOPOLOGY:  linear
MOLECULE TYPE:
       MOLECULE TYPE:  - Other nucleic Acid: synthetic
        HYPOTHETICAL:  No
          ANTI-SENSE:  No
FEATURE:
            NAME/KEY:  Primer for Polymerase Chain Reaction used to
            NAME/KEY:  make mutein of human SLPI.
PCT-US95-11445-16

Alignment Scores:
           Pred. No.:  8.76e-08         Length:      93
               Score:  146.00           Matches:     26
 Percent Similarity:   96.30%           Conservative:  0
Best Local Similarity:  96.30%           Mismatches:    1
         Query Match:  19.16%           Indels:        0
                  DB:  5                Gaps:          0

US-10-613-105-2 (1-132) x PCT-US95-11445-16  (1-93)

Qy     25 GlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSerAlaGlnCysLeu 44
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db     12 GGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCTGCCCAGTGCAAG 71

Qy     45 ArgTyrLysLysProGluCys 51
          |||||||||||||||
Db     72 AGATACAAGAAACCTGAGTGC 92

RESULT 19
US-08-379-437-3
; Sequence 3, Application US/08379437
; Patent No. 5734014
; GENERAL INFORMATION:
           APPLICANT:  ISHIMA, YOSHIAKI
           APPLICANT:  OKAWA, NORIYUKI
           APPLICANT:  YOSHIDA, MASAYA
           APPLICANT:  AMAGAYA, SAKAE
           APPLICANT:  KAJI, AKIRA
   TITLE OF INVENTION:  NOVEL ELAFIN DERIVATIVE
NUMBER OF SEQUENCES:  16
CORRESPONDENCE ADDRESS:
          ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT
             STREET:  1755 S. JEFFERSON DAVIS HWY, SUITE 400
               CITY:  ARLINGTON
              STATE:  VA
            COUNTRY:  USA
                ZIP:  22202
COMPUTER READABLE FORM:
         MEDIUM TYPE:  Floppy disk
            COMPUTER:  IBM PC compatible
    OPERATING SYSTEM:  PC-DOS/MS-DOS
            SOFTWARE:  PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER:  US/08/379,437
         FILING DATE:  27-MAR-1995
      CLASSIFICATION:  530
```

```
;      PRIOR APPLICATION DATA:
;          APPLICATION NUMBER:  JP PCT/JP93/01133
;          FILING DATE:  11-AUG-1992
;      PRIOR APPLICATION DATA:
;          APPLICATION NUMBER:  JP 234085/1992
;          FILING DATE:  11-AUG-1992
;      ATTORNEY/AGENT INFORMATION:
;          NAME:  OBLON, NORMAN F
;          REGISTRATION NUMBER:  24,618
;      TELECOMMUNICATION INFORMATION:
;          TELEPHONE:  703-413-3000
;          TELEFAX:  703-413-2220
;          TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  3:
;      SEQUENCE CHARACTERISTICS:
;          LENGTH:  177 base pairs
;          TYPE:  nucleic acid
;          STRANDEDNESS:  double
;          TOPOLOGY:  linear
;      MOLECULE TYPE:  other nucleic acid
;          DESCRIPTION:  /desc = "synthetic DNA"
;      FEATURE:
;          NAME/KEY:  CDS
;          LOCATION:  1..171
US-08-379-437-3

Alignment Scores:
Pred. No.:                    9.3e-07
Score:                        140.00
Percent Similarity:           55.36%
Best Local Similarity:        46.43%
Query Match:                  18.37%
DB:

US-10-613-105-2 (1-132) x US-08-379-437-3 (1-177)

Qy    74 AspProValAspThrProAsnProThrArgArgLysProGlyLysCysProValThrTyr 93
           ::||||||          |||         |||||||:::
Db     7 GAACCAGTTAAAGGTCCGGTGTCGACC------AAACCGGGCTCTTGCCCGATTATCCTG 60

Qy    94 GlyGlnCysLeuMetLeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysArg 113
           :::||||          |||||||          ||          |||
Db    61 ATCCGCGCCGCGTTTGCTGAACCCGCCGACCGGTTGTCTGAAAGACACTGACTGCCCGGGT 120

Qy   114 AspLeuLysCysCysMetGlyMetCysGlyLysLysSerCysValSerPro 129
           |||||||          |||          |||          |||
Db   121 ATCAAAAAATGCTGCGAAGGTTCTTGCGGTATGGCATGCTTCGTTCCG 168

RESULT 20
US-09-016-434-673
; Sequence 673, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
;      APPLICANT:  Janice Au-Young
;      APPLICANT:  Jeffrey J. Seilhamer
;      TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;      TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;      NUMBER OF SEQUENCES:  1490
;      CORRESPONDENCE ADDRESS:
;          ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;          STREET:  3174 PORTER DRIVE
;          CITY:  PALO ALTO
;          STATE:  CALIFORNIA
;          COUNTRY:  USA
;          ZIP:  94304
;      COMPUTER READABLE FORM:
;          MEDIUM TYPE:  Floppy disk
;          COMPUTER:  IBM PC compatible
;          OPERATING SYSTEM:  PC-DOS/MS-DOS
;          SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;      CURRENT APPLICATION DATA:
;          APPLICATION NUMBER:  US/09/016,434
;          FILING DATE:  HEREWITH
```

```
;      CLASSIFICATION:
;      PRIOR APPLICATION DATA:
;          APPLICATION NUMBER:
;          FILING DATE:
;      CLASSIFICATION:
;      ATTORNEY/AGENT INFORMATION:
;          NAME:  Zeller, Karen J.
;          REGISTRATION NUMBER:  37,071
;          REFERENCE/DOCKET NUMBER:  PA-0002 US
;      TELECOMMUNICATION INFORMATION:
;          TELEPHONE:  (650) 855-0555
;          TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  673:
;      SEQUENCE CHARACTERISTICS:
;          LENGTH:  234 base pairs
;          TYPE:  nucleic acid
;          STRANDEDNESS:  single
;          TOPOLOGY:  linear
;      IMMEDIATE SOURCE:
;          LIBRARY:  LUNGTUT09
;          CLONE:  2657496
US-09-016-434-673

Alignment Scores:
Pred. No.:                    1.37e-06
Score:                        140.00
Percent Similarity:           55.84%
Best Local Similarity:        41.56%
Query Match:                  18.37%
DB:

US-10-613-105-2 (1-132) x US-09-016-434-673 (1-234)

Qy     1 MetLysSerGlyLeuPheProPheLeuValLeuValAlaLeuGlyThrLeuAlaPro 20
           ::|||  |||||                ||| |||  |||
Db     5 ATGGGGTCCAGCAGCTTCTTGGTCCTCATGGTGTCTCCGTTCTTGTGACCCTGGTG--- 61

Qy    21 TrpAlaValGluGlySerGlyLysPhe---LysAlaGlyValCysProProLysLys 39
           ::|||          :::          |||
Db    62 --GCTGTGGAAGGAGTTAAAGAGGGTATAGAAGGGTTGC---CCAGCTGAC 115

Qy    40 SerAlaGlnCysLeuArgTyrLysLysProGlnSerAspTrpGlnCysProGly 59
           :::  ::|||          |||:::|||          :::
Db   116 AACGTAGCTGCTTCAGTCCGATCCTCCCCAGTGTCACACAGTCACAGGACTGTCTGGGG 175

Qy    60 LysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspProVal 76
           ::::::::|||||          |||::::
Db   176 GAAAGGAAGTGTTGTTACCTGCCACTGGCTTCAAGTGTGTGATTCCTGTG 226

RESULT 21
US-08-379-437-5
; Sequence 5, Application US/08379437
; Patent No. 5734014
; GENERAL INFORMATION:
;      APPLICANT:  ISHIMA, YOSHIAKI
;      APPLICANT:  OKAWA, NORIYUKI
;      APPLICANT:  YOSHIDA, MASAYA
;      APPLICANT:  AMAGAYA, SAKAE
;      APPLICANT:  KAJI, AKIRA
;      TITLE OF INVENTION:  NOVEL ELAFIN DERIVATIVE
;      NUMBER OF SEQUENCES:  16
;      CORRESPONDENCE ADDRESS:
;          ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT
;          STREET:  1755 S. JEFFERSON DAVIS HWY, SUITE 400
;          CITY:  ARLINGTON
;          STATE:  VA
;          COUNTRY:  USA
;          ZIP:  22202
;      COMPUTER READABLE FORM:
;          MEDIUM TYPE:  Floppy disk
;          COMPUTER:  IBM PC compatible
;          OPERATING SYSTEM:  PC-DOS/MS-DOS
;          SOFTWARE:  PatentIn Release #1.0, Version #1.30
```

```
            CURRENT APPLICATION DATA:
              APPLICATION NUMBER: US/08/379,437
              FILING DATE: 27-MAR-1995
              CLASSIFICATION: 530
            PRIOR APPLICATION DATA:
              APPLICATION NUMBER: JP PCT/JP93/01133
              FILING DATE: 11-AUG-1992
            PRIOR APPLICATION DATA:
              APPLICATION NUMBER: JP 234085/1992
              FILING DATE: 11-AUG-1992
            ATTORNEY/AGENT INFORMATION:
              NAME: OBLON, NORMAN F
              REGISTRATION NUMBER: 24,618
            TELECOMMUNICATION INFORMATION:
              TELEPHONE: 703-413-3000
              TELEFAX: 703-413-2220
              TELEX: 248855 OPAT UR
            INFORMATION FOR SEQ ID NO: 5:
            SEQUENCE CHARACTERISTICS:
              LENGTH: 177 base pairs
              TYPE: nucleic acid
              STRANDEDNESS: double
              TOPOLOGY: linear
            MOLECULE TYPE: other nucleic acid
              DESCRIPTION: /desc = "synthetic DNA"
            FEATURE:
              NAME/KEY: CDS
              LOCATION: 1..171
  US-08-379-437-5

  Alignment Scores:
  Pred. No.:                          1.19e-06      Length:              177
  Score:                              139.00        Matches:              26
  Percent Similarity:                 55.36%        Conservative:          5
  Best Local Similarity:              46.43%        Mismatches:           23
  Query Match:                        18.24%        Indels:                2
  DB:                                 1             Gaps:                  1

  US-10-613-105-2 (1-132) x US-08-379-437-5 (1-177)

  Qy    74 AspProValAspThrProAsnProThrArgArgProGlyLysCysProValThrTyr  93
             ::::||                              :::||  |::::||||||||
  Db     7 GAACCAGTTAAAGGTCGGTCGACC------AAACCGGGGCTCTTGCCCGATTATCCTG  60

  Qy    94 GlyGlnCysLeuLeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysArg 113
             :::||| |||||| :|:::|||||||||| ||||| :||||
  Db    61 ATCCGCTGCGCTGTCTGAACCCGCGAACCGTTGTCTGAAAGACACTGCCCGGGT 120

  Qy   114 AspLeuLysCysCysMetGlyMetCysGlyLysSerPro 129
             |||||  :::||| ||||||       :::|||
  Db   121 ATCAAAAATGCTGCGAAGGTTCTTGCGGTATGGCATGCTTCGTTCCG 168

  RESULT 22
  US-08-379-437-7
  ; Sequence 7, Application US/08379437
  ; Patent No. 5734014
  ; GENERAL INFORMATION:
  ;   APPLICANT: ISHIMA, YOSHIAKI
  ;   APPLICANT: OKAWA, NORIYUKI
  ;   APPLICANT: YOSHIDA, MASAYA
  ;   APPLICANT: AMAGAYA, SAKAE
  ;   APPLICANT: KAJI, AKIRA
  ;   TITLE OF INVENTION: NOVEL ELAFIN DERIVATIVE
  ;   NUMBER OF SEQUENCES: 16
  ;   CORRESPONDENCE ADDRESS:
  ;     ADDRESSEE: OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT
  ;     STREET: 1755 S. JEFFERSON DAVIS HWY, SUITE 400
  ;     CITY: ARLINGTON
  ;     STATE: VA
  ;     COUNTRY: USA
  ;     ZIP: 22202
  ;   COMPUTER READABLE FORM:
  ;     MEDIUM TYPE: Floppy disk
```

```
            COMPUTER: IBM PC compatible
            OPERATING SYSTEM: PC-DOS/MS-DOS
            SOFTWARE: Patentin Release #1.0, Version #1.30
          CURRENT APPLICATION DATA:
            APPLICATION NUMBER: US/08/379,437
            FILING DATE: 27-MAR-1995
            CLASSIFICATION: 530
          PRIOR APPLICATION DATA:
            APPLICATION NUMBER: JP PCT/JP93/01133
            FILING DATE: 11-AUG-1992
          PRIOR APPLICATION DATA:
            APPLICATION NUMBER: JP 234085/1992
            FILING DATE: 11-AUG-1992
          ATTORNEY/AGENT INFORMATION:
            NAME: OBLON, NORMAN F
            REGISTRATION NUMBER: 24,618
          TELECOMMUNICATION INFORMATION:
            TELEPHONE: 703-413-3000
            TELEFAX: 703-413-2220
            TELEX: 248855 OPAT UR
          INFORMATION FOR SEQ ID NO: 7:
          SEQUENCE CHARACTERISTICS:
            LENGTH: 177 base pairs
            TYPE: nucleic acid
            STRANDEDNESS: double
            TOPOLOGY: linear
          MOLECULE TYPE: other nucleic acid
            DESCRIPTION: /desc = "synthetic DNA"
          FEATURE:
            NAME/KEY: CDS
            LOCATION: 1..171
  US-08-379-437-7

  Alignment Scores:
  Pred. No.:                          1.19e-06      Length:              177
  Score:                              139.00        Matches:              26
  Percent Similarity:                 55.36%        Conservative:          5
  Best Local Similarity:              46.43%        Mismatches:           23
  Query Match:                        18.24%        Indels:                2
  DB:                                 1             Gaps:                  1

  US-10-613-105-2 (1-132) x US-08-379-437-7 (1-177)

  Qy    74 AspProValAspThrProAsnProThrArgArgProGlyLysCysProValThrTyr  93
             ::::||                              :::||  |::::||||||||
  Db     7 GAACCAGTTAAAGGTCGGTCGACC------AAACCGGGGCTCTTGCCCGATTATCCTG  60

  Qy    94 GlyGlnCysLeuLeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysArg 113
             :::||| |||||| :|:::|||||||||| ||||| :||||
  Db    61 ATCCGCTGCGCTGTCTGAACCCGCGAACCGTTGTCTGAAAGACACTGCCCGGGT 120

  Qy   114 AspLeuLysCysCysMetGlyMetCysGlyLysSerPro 129
             |||||  :::||| ||||||       :::|||
  Db   121 ATCAAAAATGCTGCGAAGGTTCTTGCGGTATGGCATGCTTCGTTCCG 168

  RESULT 23
  US-07-963-538B-34/c
  ; Sequence 34, Application US/07963538B
  ; Patent No. 5851983
  ; GENERAL INFORMATION:
  ;   APPLICANT: SUGIYAMA, TAKASHI
  ;   APPLICANT: KAMIMURA, TAKASHI
  ;   APPLICANT: MASUDA, KENICHI
  ;   APPLICANT: OKADA, MASAHIRO
  ;   APPLICANT: OHTSUKA, EIKO
  ;   APPLICANT: IMAIZUMI, ATSUSHI
  ;   APPLICANT: WATANABE, KUNIHITO
  ;   APPLICANT: SUGA, TETSUYA
  ;   APPLICANT: MATSUMOTO, YOHICHI
  ;   APPLICANT: TAKEUCHI, AKIKO
  ;   TITLE OF INVENTION: ELASTASE INHIBITORY POLYPEPTIDE AND
  ;   TITLE OF INVENTION: PROCESS FOR PRODUCTION THEREOF BY RECOMBINANT GENE
  ;   TITLE OF INVENTION: TECHNOLOGY
```

; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOLEY GODWARD LLP
; STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-2155
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/963,538B
; FILING DATE: 20-OCT-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/843,359
; FILING DATE: 25-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/408,483
; FILING DATE: 22-AUG-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212399
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212398
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-355553
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 62-330219
; FILING DATE: 28-DEC-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: NEELEY PH.D., RICHARD L.
; REGISTRATION NUMBER: 30,092
; REFERENCE/DOCKET NUMBER: TEJN-005/02US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-843-5070
; TELEFAX: 415-857-0663
; TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (synthetic)
US-07-963-538B-34

Alignment Scores:
Pred. No.:   1.07e-06
Score: 134.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 17.59%
DB:

Length: 68
Matches: 22
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-613-105-2 (1-132) x US-07-963-538B-34 (1-68)

Qy   98   MetLeuAsnProProAsnPheCysGluMetGlyGlnCysLysArgAspLeuLysCys   117
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66   ATGCTGAACCCGCCGAACTTCTGTGAAATGGGACGGTCAGTGTAAACGAGATCTGAAATGT   7

Qy   118   CysMet   119
        ||||||
Db   6   TGTATG   1

RESULT 24
US-09-724-864-23
; Sequence 23, Application US/09724864

; Patent No. 6380362
; GENERAL INFORMATION:
; APPLICANT: Watson, James D
; APPLICANT: Murison, James G.
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed
; TITLE OF INVENTION: by the polynucleotides and methods for their use.
; FILE REFERENCE: 11000.105OUI
; CURRENT APPLICATION NUMBER: US/09/724,864
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/171,678
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 724
; TYPE: DNA
; ORGANISM: Mouse
US-09-724-864-23

Alignment Scores:
Pred. No.:   3.24e-05
Score: 133.50
Percent Similarity: 32.20%
Best Local Similarity: 24.86%
Query Match: 17.52%
DB:

Length: 724
Matches: 44
Conservative: 13
Mismatches: 53
Indels: 67
Gaps: 9

US-10-613-105-2 (1-132) x US-09-724-864-23 (1-724)

Qy   10   LeuValLeuLeuAlaLeuGlyThrLeu------------AlaProTrpAlaVal   23
        ||| ||||||||||| ||| :::
Db   16   CTCTGGTTGCTGCGCTGGCCTCCTACTAGGGTTGCTACTGTTCACCCCCATCTCAGCC   75

Qy   24   GluGlySerGlyLysSerPheLysAlaGlyValCysProPro---LysLysSerAlaGln   42
        |:::: ||||| :::
Db   76   ACAGGCACCGATGCAGAG--AAACCCGGCGAGTGCCCCCAGCTCGAACCAATTACGGAC   132

Qy   43   CysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArg   62
        ||:::: ||||||
Db   133   TGTGTGTTG-----------GAGTGCACTTTGGACAAGGACTGTGCGGACAACCGCAAG   180

Qy   63   CysCysProAspThrCysGlyIleLysCysLeuAspPro------------   75
        ||| ||| ||| ||| |||
Db   181   TGCTGCCAGGCGGGCTGCAGCTTCGTCTGCTCCAAGCCTAATGGACCGAGCGAAGGAGAG   240

Qy   75   -------------------------------------   75

Db   241   CTCTCAGGAGCACAGATACTAAACTCTCAGAGACTGGGACTACTACTCAATCAGCGGGCCTT   300

Qy   76   -------------------ValAspThrProAsnPro------------   81
        ||| ||| |||
Db   301   GACCACTACTAAACCACCGGGAGGTCAAGTCTCCACGAAGCCACCGGCTGTGACCAGG   360

Qy   82   -----------ThrArgArgLysProGlyLysCysProValThr------------Tyr   93
        ||| ||||
Db   361   GAAGGCTAGGTGTCCGAGAAAAGCACGGCACCTGCCCAGCGTGGACATACCCAAGCTC   420

Qy   94   GlyGlnCysLeuMetLeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysArg   113
        ::: |||||||||:::|||
Db   421   GGCCTCTGT-------------GAGGACCAGTGTCAGGTGACAGCCAGTGTTCTCGC   465

Qy   114   AspLeuLysCysCysMetGlyMetCysGlyLys---SerCysValSerPro   129
        ||:::|||||||||| :::||| :::|||
Db   466   AACATGAAATGCTGCCGCCAATGGATGTGGGAAGATGGCCTGCCACCACCCC   516

RESULT 25
US-09-621-976-19119
; Sequence 19119, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.

```
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
SEQ ID NO 19119
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-19119
```

```
                                  STATE: CA
                                  COUNTRY: USA
                                  ZIP: 94306-2155
                             COMPUTER READABLE FORM:
                                  MEDIUM TYPE: Floppy disk
                                  COMPUTER: IBM PC compatible
                                  OPERATING SYSTEM: PC-DOS/MS-DOS
                                  SOFTWARE: PatentIn Release #1.0, Version #1.25
                             CURRENT APPLICATION DATA:
                                  APPLICATION NUMBER: US/07/963,538B
                                  FILING DATE: 20-OCT-1992
                                  CLASSIFICATION: 530
                             PRIOR APPLICATION DATA:
                                  APPLICATION NUMBER: US 07/843,359
                                  FILING DATE: 25-FEB-1992
                             PRIOR APPLICATION DATA:
                                  APPLICATION NUMBER: US 07/408,483
                                  FILING DATE: 22-AUG-1989
                             PRIOR APPLICATION DATA:
                                  APPLICATION NUMBER: JP 4-212399
                                  FILING DATE: 17-JUL-1992
                             PRIOR APPLICATION DATA:
                                  APPLICATION NUMBER: JP 4-212398
                                  FILING DATE: 17-JUL-1992
                             PRIOR APPLICATION DATA:
                                  APPLICATION NUMBER: JP 3-355553
                                  FILING DATE: 24-DEC-1991
                             PRIOR APPLICATION DATA:
                                  APPLICATION NUMBER: JP 62-330219
                                  FILING DATE: 28-DEC-1987
                             ATTORNEY/AGENT INFORMATION:
                                  NAME: NEELEY PH.D., RICHARD L.
                                  REGISTRATION NUMBER: 30,092
                                  REFERENCE/DOCKET NUMBER: TEJN-005/02US
                             TELECOMMUNICATION INFORMATION:
                                  TELEPHONE: 415-843-5070
                                  TELEFAX: 415-857-0663
                                  TELEX: 380816 COOLEY PA
                             INFORMATION FOR SEQ ID NO: 33:
                                  SEQUENCE CHARACTERISTICS:
                                  LENGTH: 67 base pairs
                                  TYPE: nucleic acid
                                  STRANDEDNESS: single stranded
                                  TOPOLOGY: linear
                                  MOLECULE TYPE: DNA (synthetic)
                             US-07-963-538B-33
```

```
Alignment Scores:                                      Alignment Scores:
Pred. No.:             4.43e-05                         Pred. No.:            3.55e-06
Score:                 129.50                           Score:                129.00
Percent Similarity:    34.78%                           Percent Similarity:   100.00%
Best Local Similarity: 29.71%                           Best Local Similarity: 100.00%
Query Match:           16.99%                           Query Match:          16.93%
DB:                                                     DB:
```

```
Length:  449                      Length:  67
Matches: 41                       Matches: 22
Conservative: 7                   Conservative: 0
Mismatches: 41                    Mismatches: 0
Indels: 49                        Indels: 0
Gaps: 8                           Gaps: 0
```

```
US-10-613-105-2 (1-132) x US-09-621-976-19119 (1-449)

Qy   31 LysAlaGlyValCysProProLysSerAlaGlnCysLeuArgTyrLysLysProGlu 50
           |    ||  ||:||
Db   54 AAAGAGGAGAATGCCCTCCGATAAGAAC---CCATGC-----------AAAGAGCTG 98

Qy   51 CysGlnSerAspTrpGlnProGlyLysLysArgCysLysCysProAspThrCysGlyIle 70
           ||||||| ||||          :::::::|||||||  ||||||||||||
Db   99 TGCCAGGGCGATGAATTGTCTCCGGCTGAACAGAAGTGCTGCACCACAAGCTGTGGTCGG 158

Qy   71 LysCysLeuAsp--ProValAsp-------------------------------- 77
           || ||| |||  |||
Db  159 ATCTGCCGAGACATTCCTAAGGACGGATTTCGCTCTGTTGTCGCTCAGACTAGAATGCAGTG 218

Qy   78 --------------------------------ThrProAsn 80
                                            ||
Db  219 GGCGTGATCTCGGCTCACCGCAACCTCTGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCAG 278

Qy   81 Pro---------------ThrArgArgLysProGlyLysCysPro-----ValThr 92
           ||||                |||||||            :::
Db  279 CCTCCCAAGTAGCTGGGATTTACAGGGAGGAAA---AGAGATTGCCCTAGGGTTATTCGG 335

Qy   93 TyrGlyGlnCysLeuMetLeuAsnProProAsnPheCysGluMetAspGlyGlnCysLys 112
                                    |||   |||
Db  336 AAACAATCCTGTTTG------------AAAAGGTGCATCACTGATGAGACATGTCCA 380

Qy  113 ArgAspLeuLysCysMetGlyMetCysGlyLysSerCysValSerProVal 130
                                    |||:::||
Db  381 GGTGTAAAGAWATGCTGCACGCTTGGCTGCAACAAGAGCTGTGTAGTCCCAATC 434
```

```
RESULT 26
US-07-963-538B-33
; Sequence 33, Application US/07963538B
Patent No. 5851983
; GENERAL INFORMATION:
;    APPLICANT: SUGIYAMA, TAKASHI
;    APPLICANT: KAMIMURA, TAKASHI
;    APPLICANT: MASUDA, KENICHI
;    APPLICANT: OKADA, MASAHIRO
;    APPLICANT: OHTSUKA, EIKO
;    APPLICANT: IMAIZUMI, ATSUSHI
;    APPLICANT: WATANABE, KUNIHITO
;    APPLICANT: SUGA, TETSUYA
;    APPLICANT: MATSUMOTO, YOHICHI
;    APPLICANT: TAKEUCHI, AKIKO
;    TITLE OF INVENTION: ELASTASE INHIBITORY POLYPEPTIDE AND
;    TITLE OF INVENTION: PROCESS FOR PRODUCTION THEREOF BY RECOMBINANT GENE
;    TITLE OF INVENTION: TECHNOLOGY
;    NUMBER OF SEQUENCES: 36
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE: COOLEY GODWARD LLP
;       STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
;       CITY: PALO ALTO
```

```
US-10-613-105-2 (1-132) x US-07-963-538B-33 (1-67)

Qy   95 GlnCysLeuMetLeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysArgAsp 114
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 CAGTGTCTGATGCTGAACCCGCCGAACTTCTGTGAAATGGACGGTCAGTGTAAACGAGAT 60

Qy  115 LeuLys 116
           ||||||
Db   61 CTGAAA 66
```

```
RESULT 27
US-08-304-051-7
; Sequence 7, Application US/08304051
Patent No. 5633227
; GENERAL INFORMATION:
;    APPLICANT: Muller, Daniel K.; Brownell, Elise; Delaria, Katherine A.
;    TITLE OF INVENTION: Secretory Leukocyte Protease Inhibitor as an
;    TITLE OF INVENTION: Inhibitor of Tryptase
```

```
      NUMBER OF SEQUENCES: 21
      CORRESPONDENCE ADDRESS:
         ADDRESSEE: Miles Inc.
         STREET: 400 Morgan Lane
         CITY: West Haven
         STATE: Connecticut
         COUNTRY: USA
         ZIP: 06516
      COMPUTER READABLE FORM:
         MEDIUM TYPE: Diskette, 3.50 inch, 1,300 Kb storage
         COMPUTER: Apple Macintosh
         OPERATING SYSTEM: System 7.1
         SOFTWARE: Word Perfect 3.0a
      CURRENT APPLICATION DATA:
         APPLICATION NUMBER: US/08/304,051
         FILING DATE: 12-Sept-1994
         CLASSIFICATION: 514
      PRIOR APPLICATION DATA: No. 5633227e
      ATTORNEY/AGENT INFORMATION:
         NAME: William F. Gray
         REGISTRATION NUMBER: 31018
         REFERENCE/DOCKET NUMBER: MWH 322
      TELECOMMUNICATION INFORMATION:
         TELEPHONE: (203) 937-2712
         TELEFAX: (203) 937-2795
         TELEX: 221949 MILES UR
   INFORMATION FOR SEQ ID NO: 7:
      SEQUENCE CHARACTERISTICS:
         LENGTH: 80 bases
         TYPE: nucleic acid
         STRANDEDNESS: single
         TOPOLOGY: linear
      MOLECULE TYPE:
      MOLECULE TYPE: - Other nucleic Acid: synthetic
      HYPOTHETICAL: No
      ANTI-SENSE: No
      FEATURE:
         NAME/KEY: Primer for Polymerase Chain Reaction used to make mutein
         NAME/KEY: of human SLPI.
US-08-304-051-7

Alignment Scores:
Pred. No.:                .
Score:                    4.54e-06        Length:         80
                          129.00          Matches:        22
Percent Similarity:       95.65%          Conservative:   0
Best Local Similarity:    95.65%          Mismatches:     1
Query Match:              16.93%          Indels:         0
DB:                       1               Gaps:           0

US-10-613-105-2 (1-132) x US-08-304-051-7 (1-80)

Qy  82 ThrArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsnPro 101
        |||||||||||||||||||||||||||||||||    |||||||||
Db  12 ACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTAAGATGTAGGATGCTTAACCCC 71

Qy 102 ProAsnPhe 104
        |||||||||
Db  72 CCCAATTTC 80

RESULT 28
US-08-304-051-11
; Sequence 11, Application US/08304051
; Patent No. 5633227
   GENERAL INFORMATION:
      APPLICANT: Muller, Daniel K.; Brownell, Elise; Delaria, Katherine A.
      TITLE OF INVENTION: Secretory Leukocyte Protease Inhibitor as an
      TITLE OF INVENTION: Inhibitor of Tryptase
      NUMBER OF SEQUENCES: 21
      CORRESPONDENCE ADDRESS:
         ADDRESSEE: Miles Inc.
         STREET: 400 Morgan Lane
         CITY: West Haven
         STATE: Connecticut
```

```
         COUNTRY: USA
         ZIP: 06516
      COMPUTER READABLE FORM:
         MEDIUM TYPE: Diskette, 3.50 inch, 1,300 Kb storage
         COMPUTER: Apple Macintosh
         OPERATING SYSTEM: System 7.1
         SOFTWARE: Word Perfect 3.0a
      CURRENT APPLICATION DATA:
         APPLICATION NUMBER: US/08/304,051
         FILING DATE: 12-Sept-1994
         CLASSIFICATION: 514
      PRIOR APPLICATION DATA: No. 5633227e
      ATTORNEY/AGENT INFORMATION:
         NAME: William F. Gray
         REGISTRATION NUMBER: 31018
         REFERENCE/DOCKET NUMBER: MWH 322
      TELECOMMUNICATION INFORMATION:
         TELEPHONE: (203) 937-2712
         TELEFAX: (203) 937-2795
         TELEX: 221949 MILES UR
   INFORMATION FOR SEQ ID NO: 11:
      SEQUENCE CHARACTERISTICS:
         LENGTH: 80 bases
         TYPE: nucleic acid
         STRANDEDNESS: single
         TOPOLOGY: linear
      MOLECULE TYPE:            - Other nucleic Acid: synthetic
      MOLECULE TYPE:
      HYPOTHETICAL: No
      ANTI-SENSE: No
      FEATURE:
         NAME/KEY: Primer for Polymerase Chain Reaction used to make mutein
         NAME/KEY: of human SLPI.
US-08-304-051-11

Alignment Scores:
Pred. No.:                .
Score:                    4.54e-06        Length:         80
                          129.00          Matches:        22
Percent Similarity:       95.65%          Conservative:   0
Best Local Similarity:    95.65%          Mismatches:     1
Query Match:              16.93%          Indels:         0
DB:                       1               Gaps:           0

US-10-613-105-2 (1-132) x US-08-304-051-11 (1-80)

Qy  82 ThrArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsnPro 101
        |||||||||||||||||||||||||||||||||    |||||||||
Db  12 ACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTAAGATGTAGGATGCTTAACCCC 71

Qy 102 ProAsnPhe 104
        |||||||||
Db  72 CCCAATTTC 80

RESULT 29
PCT-US95-11445-7
; Sequence 7, Application PC/TUS9511445
   GENERAL INFORMATION:
      APPLICANT: Muller, Daniel K.; Brownell, Elise; Delaria,
      APPLICANT: Katherine A.
      TITLE OF INVENTION: Secretory Leukocyte Protease
      TITLE OF INVENTION: Inhibitor as an Inhibitor of Tryptase
      NUMBER OF SEQUENCES: 21
      CORRESPONDENCE ADDRESS:
         ADDRESSEE: Bayer Corporation
         STREET: 400 Morgan Lane
         CITY: West Haven
         STATE: Connecticut
         COUNTRY: USA
         ZIP: 06516
      COMPUTER READABLE FORM:
         MEDIUM TYPE: Diskette, 3.50 inch, 1,300 Kb storage
         COMPUTER: Apple Macintosh
         OPERATING SYSTEM: System 7.1
```

```
        SOFTWARE: Word Perfect 3.0a
  CURRENT APPLICATION DATA:
     APPLICATION NUMBER: PCT/US95/11445
     FILING DATE: 11 September 1995
     CLASSIFICATION:
  PRIOR APPLICATION DATA:
     APPLICATION NUMBER: 08/304,051
     FILING DATE: 12 SEPTEMBER 1994
  ATTORNEY/AGENT INFORMATION:
     NAME: William F. Gray
     REGISTRATION NUMBER: 31018
     REFERENCE/DOCKET NUMBER: MWH 322P1
     TELEPHONE: (203) 937-2712
     TELEFAX: (203) 937-5492
INFORMATION FOR SEQ ID NO: 7:
  SEQUENCE CHARACTERISTICS:
     LENGTH: 80 bases
     TYPE: nucleic acid
     STRANDEDNESS: single
     TOPOLOGY: linear
  MOLECULE TYPE:
  MOLECULE TYPE: - Other nucleic Acid: synthetic
  HYPOTHETICAL: No
  ANTI-SENSE: No
  FEATURE:
     NAME/KEY: Primer for Polymerase Chain Reaction used to
     NAME/KEY: make mutein of human SLPI.
PCT-US95-11445-7

Alignment Scores:
Pred. No.:              4.54e-06        Length:         80
Score:                  129.00          Matches:        22
Percent Similarity:     95.65%          Conservative:    0
Best Local Similarity:  95.65%          Mismatches:      1
Query Match:            16.93%          Indels:          0
DB:                     5               Gaps:            0

US-10-613-105-2 (1-132) x PCT-US95-11445-7 (1-80)

Qy  82 ThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsnPro 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12 ACAAGGAGGAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTAAGATGCTTAACCCC 71

Qy 102 ProAsnPhe 104
          |||||||||
Db  72 CCCAATTTC 80

RESULT 30
PCT-US95-11445-11
; Sequence 11, Application PC/TUS951145
; GENERAL INFORMATION:
     APPLICANT: Muller, Daniel K.; Brownell, Elise; Delaria,
     APPLICANT: Katherine A.
     TITLE OF INVENTION: Secretory Leukocyte Protease
     TITLE OF INVENTION: Inhibitor as an Inhibitor of Tryptase
     NUMBER OF SEQUENCES: 21
  CORRESPONDENCE ADDRESS:
     ADDRESSEE: Bayer Corporation
     STREET: 400 Morgan Lane
     CITY: West Haven
     STATE: Connecticut
     COUNTRY: USA
     ZIP: 06516
  COMPUTER READABLE FORM:
     MEDIUM TYPE: Diskette, 3.50 inch, 1,300 Kb storage
     COMPUTER: Apple Macintosh
     OPERATING SYSTEM: System 7.1
     SOFTWARE: Word Perfect 3.0a
  CURRENT APPLICATION DATA:
     APPLICATION NUMBER: PCT/US95/11445
     FILING DATE: 11 September 1995
     CLASSIFICATION:
```

```
  PRIOR APPLICATION DATA:
     APPLICATION NUMBER: 08/304,051
     FILING DATE: 12 SEPTEMBER 1994
  ATTORNEY/AGENT INFORMATION:
     NAME: William F. Gray
     REGISTRATION NUMBER: 31018
     REFERENCE/DOCKET NUMBER: MWH 322P1
  TELECOMMUNICATION INFORMATION:
     TELEPHONE: (203) 937-2712
     TELEFAX: (203) 937-5492
INFORMATION FOR SEQ ID NO: 11:
  SEQUENCE CHARACTERISTICS:
     LENGTH: 80 bases
     TYPE: nucleic acid
     STRANDEDNESS: single
     TOPOLOGY: linear
  MOLECULE TYPE:
  MOLECULE TYPE: - Other nucleic Acid: synthetic
  HYPOTHETICAL: No
  ANTI-SENSE: No
  FEATURE:
     NAME/KEY: Primer for Polymerase Chain Reaction used to
     NAME/KEY: make mutein of human SLPI.
PCT-US95-11445-11

Alignment Scores:
Pred. No.:              4.54e-06        Length:         80
Score:                  129.00          Matches:        22
Percent Similarity:     95.65%          Conservative:    0
Best Local Similarity:  95.65%          Mismatches:      1
Query Match:            16.93%          Indels:          0
DB:                     5               Gaps:            0

US-10-613-105-2 (1-132) x PCT-US95-11445-11 (1-80)

Qy  82 ThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsnPro 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12 ACAAGGAGGAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTAAGATGCTTAACCCC 71

Qy 102 ProAsnPhe 104
          |||||||||
Db  72 CCCAATTTC 80

RESULT 31
US-07-963-538B-32/c
; Sequence 32, Application US/07963538B
; Patent No. 5851983
; GENERAL INFORMATION:
     APPLICANT: SUGIYAMA, TAKASHI
     APPLICANT: KAMIMURA, TAKASHI
     APPLICANT: MASUDA, KENICHI
     APPLICANT: OKADA, MASAHIRO
     APPLICANT: OHTSUKA, EIKO
     APPLICANT: IMAIZUMI, ATSUSHI
     APPLICANT: WATANABE, KUNIHITO
     APPLICANT: SUGA, TETSUYA
     APPLICANT: MATSUMOTO, YOHICHI
     APPLICANT: TAKEUCHI, AKIKO
     TITLE OF INVENTION: ELASTASE INHIBITORY POLYPEPTIDE AND
     TITLE OF INVENTION: PROCESS FOR PRODUCTION THEREOF BY RECOMBINANT GENE
     TITLE OF INVENTION: TECHNOLOGY
     NUMBER OF SEQUENCES: 36
  CORRESPONDENCE ADDRESS:
     ADDRESSEE: COOLEY GODWARD LLP
     STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
     CITY: PALO ALTO
     STATE: CA
     COUNTRY: USA
     ZIP: 94306-2155
  COMPUTER READABLE FORM:
     MEDIUM TYPE: floppy disk
     COMPUTER: IBM PC compatible
     OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;         APPLICATION NUMBER: US/07/963,538B
;         FILING DATE: 20-OCT-1992
;         CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
;         APPLICATION NUMBER: US 07/843,359
;         FILING DATE: 25-FEB-1992
; PRIOR APPLICATION DATA:
;         APPLICATION NUMBER: US 07/408,483
;         FILING DATE: 22-AUG-1989
; PRIOR APPLICATION DATA:
;         APPLICATION NUMBER: JP 4-212399
;         FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
;         APPLICATION NUMBER: JP 4-212398
;         FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
;         APPLICATION NUMBER: JP 3-355553
;         FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
;         APPLICATION NUMBER: JP 62-330219
;         FILING DATE: 28-DEC-1987
; ATTORNEY/AGENT INFORMATION:
;         NAME: NEELEY PH.D., RICHARD L.
;         REGISTRATION NUMBER: 30,092
;         REFERENCE/DOCKET NUMBER: TEJN-005/02US
; TELECOMMUNICATION INFORMATION:
;         TELEPHONE: 415-843-5070
;         TELEFAX: 415-857-0663
;         TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
;         LENGTH: 66 base pairs
;         TYPE: nucleic acid
;         STRANDEDNESS: single stranded
;         TOPOLOGY: linear
;         MOLECULE TYPE: DNA   (synthetic)
US-07-963-538B-32

Alignment Scores:
Pred. No.:                     9.25e-06          Length:            66
Score:                         125.00            Matches:           21
Percent Similarity:            100.00%           Conservative:       0
Best Local Similarity:         100.00%           Mismatches:         0
Query Match:                   16.40%            Indels:             0
DB:                            2                 Gaps:               0

US-10-613-105-2 (1-132) x US-07-963-538B-32 (1-66)

Qy  76 ValAspThrProAsnProThrArgArgProGlyLysCysProValThrTyrGlyGln 95
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  64 GTCGACACCCGAACCCGACGCGTCGTAAACCGGGTAAATGTCCGGTTACATATGGTCAG 5

Qy  96 Cys 96
        |||
Db   4 TGT 2

RESULT 32
US-07-963-538B-31
; Sequence 31, Application US/07963538B
; Patent No. 5851983
; GENERAL INFORMATION:
;         APPLICANT: SUGIYAMA, TAKASHI
;         APPLICANT: KAMIMURA, TAKASHI
;         APPLICANT: MASUDA, KENICHI
;         APPLICANT: OKADA, MASAHIRO
;         APPLICANT: OHTSUKA, EIKO
;         APPLICANT: IMAIZUMI, ATSUSHI
;         APPLICANT: WATANABE, KUNIHITO
;         APPLICANT: SUGA, TETSUYA
;         APPLICANT: MATSUMOTO, YOHICHI
;         APPLICANT: TAKEUCHI, AKIKO
```

```
;            TITLE OF INVENTION: ELASTASE INHIBITORY POLYPEPTIDE AND
;            TITLE OF INVENTION: PROCESS FOR PRODUCTION THEREOF BY RECOMBINANT GENE
;            TITLE OF INVENTION: TECHNOLOGY
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
;         ADDRESSEE: COOLEY GODWARD LLP
;         STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
;         CITY: PALO ALTO
;         STATE: CA
;         COUNTRY: USA
;         ZIP: 94306-2155
; COMPUTER READABLE FORM:
;         MEDIUM TYPE: Floppy disk
;         COMPUTER: IBM PC compatible
;         OPERATING SYSTEM: PC-DOS/MS-DOS
;         SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;         APPLICATION NUMBER: US/07/963,538B
;         FILING DATE: 20-OCT-1992
;         CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
;         APPLICATION NUMBER: US 07/843,359
;         FILING DATE: 25-FEB-1992
; PRIOR APPLICATION DATA:
;         APPLICATION NUMBER: US 07/408,483
;         FILING DATE: 22-AUG-1989
; PRIOR APPLICATION DATA:
;         APPLICATION NUMBER: JP 4-212399
;         FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
;         APPLICATION NUMBER: JP 4-212398
;         FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
;         APPLICATION NUMBER: JP 3-355553
;         FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
;         APPLICATION NUMBER: JP 62-330219
;         FILING DATE: 28-DEC-1987
; ATTORNEY/AGENT INFORMATION:
;         NAME: NEELEY PH.D., RICHARD L.
;         REGISTRATION NUMBER: 30,092
;         REFERENCE/DOCKET NUMBER: TEJN-005/02US
; TELECOMMUNICATION INFORMATION:
;         TELEPHONE: 415-843-5070
;         TELEFAX: 415-857-0663
;         TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
;         LENGTH: 63 base pairs
;         TYPE: nucleic acid
;         STRANDEDNESS: single stranded
;         TOPOLOGY: linear
;         MOLECULE TYPE: DNA   (synthetic)
US-07-963-538B-31

Alignment Scores:
Pred. No.:                     1.11e-05          Length:            63
Score:                         124.00            Matches:           21
Percent Similarity:            100.00%           Conservative:       0
Best Local Similarity:         100.00%           Mismatches:         0
Query Match:                   16.27%            Indels:             0
DB:                            2                 Gaps:               0

US-10-613-105-2 (1-132) x US-07-963-538B-31 (1-63)

Qy  74 AspProValAspThrProAsnProThrArgArgProGlyLysCysProValThrTyr 93
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GATCCGGTCGACACCCGAACCCGACGCGTCGTAAACCGGGTAAATGTCCGGTTACATAT 60

Qy  94 Gly 94
        |||
Db  61 GGT 63
```

RESULT 33
US-08-304-051-12
; Sequence 12, Application US/08304051
; Patent No. 5633227
; GENERAL INFORMATION:
; APPLICANT: Muller, Daniel K.; Brownell, Elise; Delaria, Katherine A.
; TITLE OF INVENTION: Secretory Leukocyte Protease Inhibitor as an
; TITLE OF INVENTION: Inhibitor of Tryptase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Miles Inc.
; STREET: 400 Morgan Lane
; CITY: West Haven
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06516
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1,300 Kb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word Perfect 3.0a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/304,051
; FILING DATE: 12-Sept-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA: No. 5633227e
; ATTORNEY/AGENT INFORMATION:
; NAME: William F. Gray
; REGISTRATION NUMBER: 31018
; REFERENCE/DOCKET NUMBER: MWH 322
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 937-2712
; TELEFAX: (203) 937-2795
; TELEX: 221949 MILES UR
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 89 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; MOLECULE TYPE: - Other nucleic Acid: synthetic
; HYPOTHETICAL: No
; ANTI-SENSE: No
; FEATURE:
; NAME/KEY: Primer for Polymerase Chain Reaction used to make mutein
; NAME/KEY: of human SLPI.
US-08-304-051-12

Alignment Scores:
Pred. No.:
Score: 124.00
Percent Similarity: 92.86%
Best Local Similarity: 92.86%
Query Match: 16.27%
DB:

Length: 89
Matches: 26
Conservative: 0
Mismatches: 2
Indels: 2
Gaps: 1

US-10-613-105-2 (1-132) x US-08-304-051-12 (1-89)

Qy 25 GlySerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCys--- 43
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   9 GGCTCTGGAAAGTCCTTCAAAGCTGGAGTCGT---CCTAAGAAATCTGCCCAGTGCAAG 65

Qy 44 LeuArgTyrLysLysProGluCys 51
           ||||||||||||||||
Db 66 CTTAGATACAAGAAACCTGAGTGC 89

RESULT 34
US-08-304-051-13
; Sequence 13, Application US/08304051
; Patent No. 563227
; GENERAL INFORMATION:
; APPLICANT: Muller, Daniel K.; Brownell, Elise; Delaria, Katherine A.

TITLE OF INVENTION: Secretory Leukocyte Protease Inhibitor as an
; TITLE OF INVENTION: Inhibitor of Tryptase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Miles Inc.
; STREET: 400 Morgan Lane
; CITY: West Haven
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06516
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1,300 Kb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word Perfect 3.0a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/304,051
; FILING DATE: 12-Sept-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA: No. 5633227e
; ATTORNEY/AGENT INFORMATION:
; NAME: William F. Gray
; REGISTRATION NUMBER: 31018
; REFERENCE/DOCKET NUMBER: MWH 322
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 937-2712
; TELEFAX: (203) 937-2795
; TELEX: 221949 MILES UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 89 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; MOLECULE TYPE: - Other nucleic Acid: synthetic
; HYPOTHETICAL: No
; ANTI-SENSE: No
; FEATURE:
; NAME/KEY: Primer for Polymerase Chain Reaction used to make mutein
; NAME/KEY: of human SLPI.
US-08-304-051-13

Alignment Scores:
Pred. No.:
Score: 124.00
Percent Similarity: 92.86%
Best Local Similarity: 92.86%
Query Match: 16.27%
DB:

Length: 89
Matches: 26
Conservative: 0
Mismatches: 2
Indels: 2
Gaps: 1

US-10-613-105-2 (1-132) x US-08-304-051-13 (1-89)

Qy 25 GlySerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCys--- 43
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   9 GGCTCTGGAAAGTCCTTCAAAGCTGGAGTCGT---CCTAAGAAATCTGCCCAGTGCAAG 65

Qy 44 LeuArgTyrLysLysProGluCys 51
           ||||||||||||||||
Db 66 CTTAGATACAAGAAACCTGAGTGC 89

RESULT 35
PCT-US95-11445-12
; Sequence 12, Application PC/TUS9511445
; GENERAL INFORMATION:
; APPLICANT: Muller, Daniel K.; Brownell, Elise, Delaria,
; APPLICANT: Katherine A.
; TITLE OF INVENTION: Secretory Leukocyte Protease
; TITLE OF INVENTION: Inhibitor as an Inhibitor of Tryptase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bayer Corporation
; STREET: 400 Morgan Lane

CITY: West Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06516
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1,300 Kb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word Perfect 3.0a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/11445
FILING DATE: 11 September 1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/304,051
FILING DATE: 12 SEPTEMBER 1994
ATTORNEY/AGENT INFORMATION:
NAME: William F. Gray
REGISTRATION NUMBER: 31018
REFERENCE/DOCKET NUMBER: MWH 322P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 937-2712
TELEFAX: (203) 937-5492
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 89 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: - Other nucleic Acid: synthetic
HYPOTHETICAL: No
ANTI-SENSE: No
FEATURE:
NAME/KEY: Primer for Polymerase Chain Reaction used to
NAME/KEY: make mutein of human SLPI.
PCT-US95-11445-12

Alignment Scores:
Pred. No.: 1.79e-05
Score: 124.00
Percent Similarity: 92.86%
Best Local Similarity: 92.86%
Query Match: 16.27%
DB: 5

Length: 89
Matches: 26
Conservative: 0
Mismatches: 0
Indels: 2
Gaps: 2

US-10-613-105-2 (1-132) x PCT-US95-11445-12 (1-89)

Qy 25 GlySerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCys--- 43
       ||||||||||||||||||||||||||||||||||||||||||||
Db   9 GGCTCTGGAAAGTCCTTCAAAGCTGGAGTCGT---CCTAAGAAATCTGCCCAGTGCAAG 65

Qy 44 LeuArgTyrLysLysProGluCys 51
       ||||||||||||||||||||||||
Db 66 CTTAGATACAAGAAACCTGAGTGC 89

RESULT 36
PCT-US95-11445-13
; Sequence 13, Application PC/TUS9511445
; GENERAL INFORMATION:
; APPLICANT: Muller, Daniel K.; Brownell, Elise; Delaria,
; APPLICANT: Katherine A.
; TITLE OF INVENTION: Secretory Leukocyte Protease
; TITLE OF INVENTION: Inhibitor as an Inhibitor of Tryptase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bayer Corporation
; STREET: 400 Morgan Lane
; CITY: West Haven
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06516
; COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1,300 Kb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word Perfect 3.0a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/11445
FILING DATE: 11 September 1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/304,051
FILING DATE: 12 SEPTEMBER 1994
ATTORNEY/AGENT INFORMATION:
NAME: William F. Gray
REGISTRATION NUMBER: 31018
REFERENCE/DOCKET NUMBER: MWH 322P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 937-2712
TELEFAX: (203) 937-5492
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 89 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: - Other nucleic Acid: synthetic
HYPOTHETICAL: No
ANTI-SENSE: No
FEATURE:
NAME/KEY: Primer for Polymerase Chain Reaction used to
NAME/KEY: make mutein of human SLPI.
PCT-US95-11445-13

Alignment Scores:
Pred. No.: 1.79e-05
Score: 124.00
Percent Similarity: 92.86%
Best Local Similarity: 92.86%
Query Match: 16.27%
DB: 5

Length: 89
Matches: 26
Conservative: 0
Mismatches: 0
Indels: 2
Gaps: 2

US-10-613-105-2 (1-132) x PCT-US95-11445-13 (1-89)

Qy 25 GlySerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCys--- 43
       ||||||||||||||||||||||||||||||||||||||||||||
Db   9 GGCTCTGGAAAGTCCTTCAAAGCTGGAGTCGT---CCTAAGAAATCTGCCCAGTGCAAG 65

Qy 44 LeuArgTyrLysLysProGluCys 51
       ||||||||||||||||||||||||
Db 66 CTTAGATACAAGAAACCTGAGTGC 89

RESULT 37
US-09-152-060-34
; Sequence 34, Application US/09152060
; Patent No. 6448230
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: PZ003P1.US
; CURRENT APPLICATION NUMBER: US/09/152,060
; CURRENT FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: PCT/US98/04858
; EARLIER FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 60/040,762
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/040,710
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/050,934
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,100
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,357
; EARLIER FILING DATE: 1997-05-30

; EARLIER APPLICATION NUMBER: 60/048,189
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/057,765
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/068,368
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-152-060-34

Alignment Scores:
Pred. No.:   0.00171
Score: 117.50
Percent Similarity: 38.12%
Best Local Similarity: 29.38%
Query Match: 15.42%
DB: 4

Length: 753
Matches: 47
Conservative: 14
Mismatches: 58
Indels: 41
Gaps: 10

US-10-613-105-2 (1-132) x US-09-152-060-34 (1-753)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro 20
             ||| |||||
Db   31 ATGGGATCTTCTGGACTTTTGAGCCTCCTGTGTCTATTCGTTCCTC-----TTAGCG--- 81

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysSerAlaGlyValCysPro-------Pro 37
             || ||::||||| ::: |||||
Db   82 ---AATGTCCAGGGACCTGGT------CTGACTGATTGGTTATTTCCCAGGAGATGTCCC 132

Qy   38 LysLysSerAlaGlnCysLeuArgTyrLysLeuProGluCysGlnSerAspTrpGlnCys 57
             |||||||::::::
Db 133 AAAATCAGAGAGAAGAATGTGAATTCCAGAAAGGGATGTGTACAAAGGACAGACAATGC 192

Qy   58 ProGlyLysLysArgCysCysProAspThrCysGlyIleLysCysLeu----------- 73
             ||:||||||| ::||||||
Db 193 CAGGACAACAAGAGTGTTGTGTCTTCAGCTGCGGAAAAAAATGTTTAGATCTCAAACAA 252

Qy   74 AspProValAspThrProAsnProThr-------------------------------- 82
             ||| ::: ||
Db 253 GATGTATGCGAAATGCCAAAAGAAACTGGCCCCCGGCTTATTTCTTCATTGGTGG 312

Qy   83 --ArgArgLysProGlyLysCysProVal---ThrTyrGlyLysCys----------- 96
             :::||||||| :::
Db 313 TATGACAAGAAAGATATAACTTGCTCCATGTTGTCTATGGTGGCTGCCAGGGAAACAAT 372

Qy   97 ------------LeuMetLeuAsnProProAsnPheCysGlu------------------ 106
                        |||:::
Db 373 AACAACTTCAATCCAAAGCCAACTGCCTGAACACCTGCAAGAATAAACGCTTTCCCTGA 432

Qy 107 MetAspGlyGlnCysLysArgAspLeuCysCysMetGlyMetCysGlyLysSerCys 126
             :::|||| ||| :::
Db 433 TTGGATAAGGATGCACTGGAAGAACTGCCAGAATGT---GGCTCATGTCTGAGTACTGT 489

RESULT 38
US-09-270-767-1693
; Sequence 1693, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1693
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1693

Alignment Scores:
Pred. No.:   0.00304
Score: 112.50
Percent Similarity: 43.42%
Best Local Similarity: 32.89%
Query Match: 14.76%
DB: 4

Length: 472
Matches: 25
Conservative: 8
Mismatches: 18
Indels: 25
Gaps: 3

US-10-613-105-2 (1-132) x US-09-270-767-1693 (1-472)

Qy   25 GlySerGlyLys------------------------------SerPheLysAlaGlyValCysPro-- 36
             ||||||||| ||| :::
Db 220 GGCTCTCGAAAGGAATGCCGAGTGGTGGACGTGGCGAGGGTGAATACTGTCCTCCA 279

Qy   37 ------------ProLysLysSerAlaGlnCysLeuArgTyrLysLeuLysPro------- 49
                        |||:::||
Db 280 GTCCCCGCTGTTGCCTAGGAAGCCGGGACAGTGTCCTTACCTGGTTCCACCGGGTCCT 339

Qy   50 ------------GluCysGlnSerAspTrpGlnCysProGly 59
                        ||||||:::|||
Db 340 GATAACCTGGATGCCAACACCTGTCCTACCGAGTGCCTACCGATGCCCATTGTGATGGA 399

Qy   60 LysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspPro 75
             :::||||||||| ::||||:::|||||
Db 400 GCCCGACGCGTGTTGCTCCAATCGATGTGGCACCCAGTGTGTGGATCCT 447

RESULT 39
US-09-270-767-16975
; Sequence 16975, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16975
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-16975

Alignment Scores:
Pred. No.:   0.00304
Score: 112.50
Percent Similarity: 43.42%
Best Local Similarity: 32.89%
Query Match: 14.76%
DB: 4

Length: 472
Matches: 25
Conservative: 8
Mismatches: 18
Indels: 25
Gaps: 3

US-10-613-105-2 (1-132) x US-09-270-767-16975 (1-472)

Qy   25 GlySerGlyLys------------------------------SerPheLysAlaGlyValCysPro-- 36
             ||||||||| ||| :::
Db 220 GGCTCTCGAAAGGAATGCCGAGTGGTGGACGTGGCGAGGGTGAATACTGTCCTCCA 279

Qy   37 ------------ProLysLysSerAlaGlnCysLeuArgTyrLysLeuLysPro------- 49
                        |||:::||
Db 280 GTCCCCGCTGTTGCCTAGGAAGCCGGGACAGTGTCCTTACCTGGTTCCACCGGGTCCT 339

Qy   50 ------------GluCysGlnSerAspTrpGlnCysProGly 59
                        ||||||:::|||
Db 340 GATAACCTGGATGCCAACACCTGTCCTACCGAGTGCCTACCGATGCCCATTGTGATGGA 399

Qy   60 LysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspPro 75
             :::||||||||| ::||||:::|||||
Db 400 GCCCGACGCGTGTTGCTCCAATCGATGTGGCACCCAGTGTGTGGATCCT 447

RESULT 40
US-09-152-060-50
; Sequence 50, Application US/09152060
; Patent No. 6448230
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: PZ003P1.US
; CURRENT APPLICATION NUMBER: US/09/152,060
; CURRENT FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: PCT/US98/04858
; EARLIER APPLICATION NUMBER: 60/040,762
; EARLIER FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 60/040,710
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/050,934
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/048,100
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,357
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,189
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/057,765
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/068,368
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-152-060-50

Alignment Scores:
Pred. No.:   0.00614   Length: 783
Score: 112.50   Matches: 47
Percent Similarity: 38.51%   Conservative: 15
Best Local Similarity: 29.19%   Mismatches: 57
Query Match: 14.76%   Indels: 42
                                           Gaps: 10
DB:

US-10-613-105-2 (1-132) x US-09-152-060-50 (1-783)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuAlaGlyThrLeuAlaPro 20
            ||| ||||| || ||||
Db   31 ATGGGMTCTTCTGGACTTTTGAGCCTCTGTGCTATTCGTCCTC------TTAGCG---- 81

Qy   21 TrpAlaValGluGluGlyLysSerGlyPheLysSerAlaGlyValCysPro----------Pro 37
            |||:::|| ||||||| :::   ::
Db   82 ---AATGTCCAGGAGACCTGT------CTGACTGATTGGTTATTTCCCAGGAGATGTCCC 132

Qy   38 LysLysSerAlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCys 57
            :::::| ::::: |||||:
Db 133 AAAATCAGAGAAGAGATGTGAATTCCAAGAAGGAGATGTGTGTACAAAGGACAGACAATGC 192

Qy   58 ProGlyLysLysArgCysCysProAspThrCysGlyIleLysCysLeu----------- 73
            ||:||||||| :::||||||
Db 193 CAGGACAACAAGAAGTGTTGTGTCTTCAGCTGGCGAAAAAAATGTTTAGATCTCAAACAA 252

Qy   74 AspProValAspThrProAsnProThr----------------------- 82
            || :::||||||
Db 253 GATGTATGCGAAAATGCCAAAAAGAAAACTGGCCCCTGCCTTATTTCTTCATTCGTGG 312

Qy   83 --ArgArgLysProGlyLysLys---ThrTyrGlyGlnCys----------- 96
            :::||||||| ||||:||
Db 313 TATGACAAGAAAGAAGATAATAACTTGCTCCATGTTTGTCTCATGCCTCCATGTCTGGTGGCTATGCTGGGGAAACAA 372

Qy   97 -----------LeuMetLeuAsnProPro-AsnPheCysGlu- 106
                                     ||||| |||||||
Db 373 TAACACTTCCAATCCAAAGCCAACTGCCTGAACACCTGCAAGAATAAACGCTTCCCTG 432

Qy 107 -MetAspGlyGlnCysArgLysArgAspLeuLysCysMetCysGlyLysLysSerCy 126
           :::|| ||| ||||| :::
Db 433 ATTGGATAAGGATGCACTGGAGAGACTGCCAGAATGT---GGCTCATGTGCTCTGAGTACTG 489

Qy 126 s 126

Db 490 T 490

RESULT 41
US-09-383-586-29
; Sequence 29, Application US/09383586
; Patent No. 622419
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Abernethy, Nevin
; APPLICANT: Onrust, Rene
; APPLICANT: Kumble, Anand
; APPLICANT: Murison, Greg
; TITLE OF INVENTION: Compounds isolated from stromal cells
; TITLE OF INVENTION: and methods for their use
; FILE REFERENCE: 11000.1037c1
; CURRENT APPLICATION NUMBER: US/09/383,586
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 437
; TYPE: DNA
; ORGANISM: Mouse
US-09-383-586-29

Alignment Scores:
Pred. No.:   0.00568   Length: 437
Score: 109.50   Matches: 39
Percent Similarity: 39.71%   Conservative: 15
Best Local Similarity: 28.68%   Mismatches: 61
Query Match: 14.37%   Indels: 21
                                           Gaps: 7
DB:

US-10-613-105-2 (1-132) x US-09-383-586-29 (1-437)

Qy   9 PheLeuVal---LeuLeuAlaLeuGlyThrLeuAlaProAlaValAlaGluGlySerGly 27
            |||||||| ::: ||
Db   11 TTTCTGGTGGCTTGATCACTGTGGGGATGAACACTACCTATGTGTCTTCGCCCCAAA 70

Qy   28 LysSerPheLysAlaGlyValCysProProLysSerProLysSerAlaGlnCysLeuArgTyrLys 47
            ::   :::
Db   71 GAATTTGAAAAACCTGGAGCTTGT---CCCAAGCCTTCACCAGAAAGTGTTGGAATTTGT 127

Qy   48 LysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysProAspThr 67
            :::|||| ||||||| :::|||
Db 128 GTTGATCAATGCTCAGGAGATGGATCCGTCCCTGGCAACATGAAGTGCTGTAGCAATAGC 187

Qy   68 CysGlyIleLysCysLeuAspProVal-----------------------AspThr 78
            ||| :::
Db 188 TGTGGTCATGTCTGCAAAACCTCGTCTTTAAATGGTTGACAGCCATGTGGAAGATGA 247

Qy   79 ProAsnProThrArgArgLysProGlyLysLysCysProValThrTyr----------Gly 94
            ||| ||| ||| :::|||
Db 248 TTCAATCTTCATAAACATGAATGATGGCCAGCCCAGAAGATTTCTTCTGAATTCACAGA 307

Qy   95 GlnCysLeuMetLeuAsnPro---ProAsnPheCysGluMetCysGlyLysLysArg 113
            :::||||| ||| :::|||
Db 308 GCCTGTCGTTGGCTACTTCCTAGCCCTAGAATTGCATTCTTGGAC-------AAGGAA 358

Qy 114 AspLeuLysCysCysMetCysGlyLysLysSerCysValSerPro 129
            ::: ||| :::
Db 359 GATCTATATTGT-----GGTGACAATGCCCTAATATGTCTGTGTCCA 400

RESULT 42
US-09-823-038A-29
; Sequence 29, Application US/09823038A
; Patent No. 6797271
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Abernethy, Nevin
; APPLICANT: Onrust, Rene
; APPLICANT: Kumble, Anand
; APPLICANT: Murison, Greg
; TITLE OF INVENTION: Compositions Isolated From Stromal Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1037c3
; CURRENT APPLICATION NUMBER: US/09/823,038A
; CURRENT FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 437
; TYPE: DNA
; ORGANISM: Mouse
US-09-823-038A-29

Alignment Scores:
Pred. No.:   0.00568
Score: 109.50   Length: 437
Percent Similarity: 39.71%   Matches: 39
Best Local Similarity: 28.68%   Conservative: 15
Query Match: 14.37%   Mismatches: 61
DB: Indels: 21
                                        Gaps: 7

US-10-613-105-2 (1-132) x US-09-823-038A-29 (1-437)

Qy   9 PheLeuVal---LeuLeuAlaLeuGlyThrLeuAlaProTrpAlaValGluGlySerGly   27
             :::||| :::|||::::||||
Db   11 TTTCTGGTGGCTTTGATCACTGTGGGGATGAACACTACCTATGTGTCTTGCCCCAAA   70

Qy   28 LysSerPheLysAlaGlyValCysProGlyLysSerAlaGlnCysLeuArgTyrLys   47
             ::: ||||||| ::: ||| ::: |||
Db   71 GAATTTGAAAAACCTGGAGCTTGT---CCCAAGCCTCACCAGAAGTGTTGGAATTTGT   127

Qy   48 LysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysProAspThr   67
             ::||| ||||||| ::: |||||| :::::::
Db   128 GTTGATCAATGCTCAGGAGATGGATCCTGCCCTGGCAACATGAAGTGCTGTAGCAATAGC   187

Qy   68 CysGlyIleLysCysLeuAspProVal---------------------------AspThr   78
             |||| |||
Db   188 TGTGGTCATGTCTGCAAAACTCCTGTCTTTTAAATGGTTGACAGCCATGTGGAAGATGGA   247

Qy   79 ProAsnProThrArgLysProGlyLysCysProValThrTyr--------------Gly   94
             || :::
Db   248 TTCAATCTTCATAAACATGATGATGGCCAGCCCAGAAGATTTCTTCTGAATTCACAGA   307

Qy   95 GlnCysLeuMetLeuAsnPro---ProAsnPheCysGluMetAspGlyGlnCysLysArg 113
             :::||| ||| |||||||||||
Db   308 GCCTGTGGCTTGGCTACTTCCTAGCCTAGACCCTTAGAATTGCATTCTTGGAC------- 358

Qy   114 AspLeuLysCysCysMetGlyMetCysGlyLysLysSerCysValSerPro   129
             :|||||||| ::||||||
Db   359 GATCTATATTGT------GGTGACAATGCCCAATAATGTCTGTGTCCA   400

RESULT 43
US-07-978-895-3
; Sequence 3, Application US/07978895
; Patent No. 5480968
; GENERAL INFORMATION:
; APPLICANT: Kraus, Matthias H.
; APPLICANT: Aaronson, Stuart A.
; TITLE OF INVENTION: AN ISOLATED POLYPEPTIDE RELATED TO THE
; TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR RECEPTOR, ANTIGEN THERETO, AND
; TITLE OF INVENTION: BIOASSAYS AND METHODS RELATED THERETO
; NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:
; ADDRESSEE: Suite 400
; STREET: 133 Carnegie Way, N.W.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.A.
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/978,895
; FILING DATE: 19921110
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/444,406
; FILING DATE: 01-DEC-1989
; REGISTRATION NUMBER: 33,438
; ATTORNEY/AGENT INFORMATION:
; NAME: Perryman, David G.
; REFERENCE/DOCKET NUMBER: 1414-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 688-0770
; TELEFAX: (404) 688-9880
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4905 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 100..4125
US-07-978-895-3

Alignment Scores:
Pred. No.:   0.186
Score: 109.00   Length: 4905
Percent Similarity: 32.32%   Matches: 41
Best Local Similarity: 25.00%   Conservative: 12
Query Match: 14.30%   Mismatches: 43
DB: Indels: 68
                                        Gaps: 9

US-10-613-105-2 (1-132) x US-07-978-895-3 (1-4905)

Qy   23 ValGluGlySerGlyLysSerPheLysAlaGlyValCysProPro-------------   37
             |||::: |||:::|||:::||||
Db   625 GTGAAGGACAATGGCAGAAGC----------------TGTCCCCCGTGTCATGAGGTTTGC 669

Qy   38 -----------------------LysLysSerAlaGlnCysLeuArgTyrLysLys-----   48
                                      |||
Db   670 AAGGGGGATGCTGGGGTCCTGGATCAGAAGACTGCCAGACATTGACCAAGACCATCTGT 729

Qy   49 ---ProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysProAspThr   67
               |||||:::||| ::|||||
Db   730 GCTCCTCAGTGTAATGTCACTGCTTTGGGCCCAACCCCACCAGTGCTGCCATGATGAG 789

Qy   68 CysGlyIleLysCysLeuAspProValAspThr--------------------------   78
             ||| |||||
Db   790 TGTGCCGGGGCGTCGCTCAGGCCTCCAGGACACAGACTGCTTGCCTGCCGGCACTTCAAT 849

Qy   78 --   78
Db   850 GACAGTGGAGCCTGTGTACCTCGCTGTCTGCTCCAGCGCTCTTGTCTACACAACAAGCTAACTTTC 909

Qy   79 ---ProAsnPro---ThrArgArgLysProGly------------------------LysCys   89
               ||||||||| ::: ::: |||
Db   910 CAGCTGAACCAATCCCAACCCACACCAAGTATCAGTATGGAGAGTTTGTAGCCAGCTGT 969

Qy   90 Pro-------------ValThrTyrGlyGlnCysLeuMetLeuAsnProProAsnPheCys 105

```
Db   970 CCCCATAACTTGTGTGGTGGATCAACATCTGTGTCAGGGCCTGTCCTCCTGACAAGATG 1029
             |||||                    |||||:::
Qy   106 GluMetAspGlyGlnCysLysArgAspLeuLysCysMet--------GlyMetCys 122
             |||:::|||                         |||  |||:::|||
Db  1030 GAAGTAGAT--------AAAAATGGGCTCAAGATGTGTGAGCCTTGTGGGGACTATGT 1080

Qy   123 GlyLysSerCys 126
             |||:::|:||
Db  1081 CCCAAGCCTGT 1092


RESULT 44
US-08-473-119-3
; Sequence 3, Application US/08473119
; Patent No. 5820859
; GENERAL INFORMATION:
;   APPLICANT: Kraus, Matthias H.
;   APPLICANT: Aaronson, Stuart A.
;   TITLE OF INVENTION: AN ISOLATED POLYPEPTIDE RELATED TO THE
;   TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR RECEPTOR, ANTIGEN THERETO, AND
;   TITLE OF INVENTION: BIOASSAYS AND METHODS RELATED THERETO
;   NUMBER OF SEQUENCES: 12
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Suite 400
;     STREET: 133 Carnegie Way, N.W.
;     CITY: Atlanta
;     STATE: Georgia
;     COUNTRY: U.S.A.
;     ZIP: 30303
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/473,119
;     FILING DATE: 07-JUN-1995
;     CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 07/978,895
;     FILING DATE: 10-NOV-1992
;     APPLICATION NUMBER: US 07/444,406
;     FILING DATE: 01-DEC-1989
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Perryman, David G.
;     REGISTRATION NUMBER: 33,438
;     REFERENCE/DOCKET NUMBER: 1414-028
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (404) 688-0770
;     TELEFAX: (404) 688-9880
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 4905 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 100..4125
US-08-473-119-3

Alignment Scores:
Pred. No.:                0.186        Length:          4905
Score:                  109.00        Matches:           41
Percent Similarity:      32.32%       Conservative:      12
Best Local Similarity:   25.00%       Mismatches:        43
Query Match:             14.30%       Indels:            68
DB:                                   Gaps:               9

US-10-613-105-2 (1-132) x US-08-473-119-3 (1-4905)

Qy    23 ValGluGlySerPheLysSerPheLysAlaGlyValCysProPro----------- 37
```

```
Db   625 GTGAAGGACAATGGCAGAAGC--- 
             |||:::       :::|||:||||
Qy    38 ----------LysLysSerAlaGlnCysLeuArgTyrLysLys 48
                          |||           |||
Db   670 AAGGGGCGATGCTGGGGTCCTGATCAGAAGACTGCCAGACATTGACCAGACCATCTGT 729

Qy    49 ---ProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysProAspThr 67
             ::|||:::|||                         ::|||:::|||
Db   730 GCTCCTCCAGTGTAATGTCACTGCTTTGGGCCCAACCCCAACCAGTGTGCCATGATGAG 789

Qy    68 CysGlyIleLysCysLeuAspProValAspThr---------- 78
             |||       |||||||
Db   790 TGTGCCGGGGCTGTCTCAGGCCCTCAGGACACAGACTGCTTTGCCTGCCGGGCACTTCAAT 849

Qy    78 ----------------------------------- 78
Db   850 GACAGTGGAGCCTGTACCTGCTGTCCACAGCCCTCTTGTCTACACAAGAGCTAACTTTC 909

Qy    79 -------ProAsnPro---ThrArgArgProGly----------LysCys 89
                        |||:||:::       |||   ::: :::|||
Db   910 CAGCTGGAACCCAATCCCACACCAAGTATCAGTATGGAGAGTTTGTAGCCAGCTGT 969

Qy    90 Pro----------------ValThrTyrGlyGlnCysLeuMetLeuAsnProProAsnPheCys 105
             |||                |||  |||:::
Db   970 CCCCATAACTTTGTGGTGGATCAAACATCCTGTGTCAGGGCCTGTCCTCCTGACAAGATG 1029

Qy   106 GluMetAspGlyGlnCysLysArgAspLeuLysCysMet--------GlyMetCys 122
             |||:::|||                         |||  |||:::|||
Db  1030 GAAGTAGAT--------AAAAATGGGCTCAAGATGTGTGAGCCTTGTGGGGACTATGT 1080

Qy   123 GlyLysSerCys 126
             |||:::|:||
Db  1081 CCCAAGCCTGT 1092


RESULT 45
US-08-475-352-3
; Sequence 3, Application US/08475352
; Patent No. 5916755
; GENERAL INFORMATION:
;   APPLICANT: Kraus, Matthias H.
;   APPLICANT: Aaronson, Stuart A.
;   TITLE OF INVENTION: AN ISOLATED POLYPEPTIDE RELATED TO THE
;   TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR RECEPTOR, ANTIGEN THERETO, AND
;   TITLE OF INVENTION: BIOASSAYS AND METHODS RELATED THERETO
;   NUMBER OF SEQUENCES: 12
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Suite 400
;     STREET: 133 Carnegie Way, N.W.
;     CITY: Atlanta
;     STATE: Georgia
;     COUNTRY: U.S.A.
;     ZIP: 30303
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/475,352
;     FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 07/978,895
;     FILING DATE:
;     APPLICATION NUMBER: US 07/444,406
;     FILING DATE: 01-DEC-1989
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Perryman, David G.
;     REGISTRATION NUMBER: 33,438
;     REFERENCE/DOCKET NUMBER: 1414-028
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (404) 688-0770
```

```
;        TELEFAX:    (404) 688-9980
; INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;          LENGTH: 4905 base pairs
;          TYPE: nucleic acid
;          STRANDEDNESS: single
;          TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        FEATURE:
;          NAME/KEY:  CDS
;          LOCATION:  100..4125
US-08-475-352-3

Alignment Scores:
Pred. No.:               0.186            Length:          4905
Score:                   109.00           Matches:         41
Percent Similarity:      32.32%           Conservative:    12
Best Local Similarity:   25.00%           Mismatches:      43
Query Match:             14.30%           Indels:          68
DB:                      2                Gaps:            9

US-10-613-105-2 (1-132) x US-08-475-352-3 (1-4905)

Qy    23 ValGluGlySerGlyLysPheLysAlaGlyValCysProPro----------------  37
            |||::|||||| :::|||||::::|||
Db   625 GTGAAGGACAATGGCAGAAGC-------------TGTCCCCCCTGTCATGAGGTTTGC  669

Qy    38 ----------------LysLysSerAlaGlnCysLeuArgTyrLysLys----------  48
                          ||| |||
Db   670 AAGGGGCGATGCTGGGGTCCTGGATCAGAGAGACTGCCAGACATTGACCAAGACCATCTGT  729

Qy    49 ---ProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysProAspThr  67
            |||::||| |||           ::||||||||
Db   730 GCTCCTCCAGTGTAATGGTCACTGCTTTGGGCCCAACCCCAACCAGTGCTGCCATGATGAG  789

Qy    68 CysGlyIleLysCysLeuAspProValAspThr----------------------  78
            ||| ||||||
Db   790 TGTGCCGGGGGCTGCTCAGGCCCTCAGGACACCAGGACTGCTTTGCCTGCCGGCACTTCAAT  849

Qy    78 --------------------------------------------------------  78
Db   850 GACAGTGGAGCCGCTGTACCTCGCTCGTCCACCAGCCTCTTGCTCACAACAGCTAACTTTC  909

Qy    79 --------ProAsnPro---ThrArgLysProGly-----------LysCys  89
                    |||||:||| |||:::||          |||
Db   910 CAGCTGGAACCCAATCCCACCAGTATCAGTATGGAGGAGTTTGTGTAGCCAGCTGT  969

Qy    90 Pro---------ValThrTyrGlyGlnCysLeuMetLeuAsnProProAsnPheCys  105
            ||:::         |||:::          |||||::::
Db   970 CCCCATAAACTTTGTGTGGATCAAACATCCTGTGTCAGGGCGTGTCCTCCTGACAAGATG  1029

Qy   106 GluMetAspGlyGlyGlnCysLysArgAspLeuLysCysCysMet-------GlyMetCys  122
            ||:::|||||           ||||||            |||    ||| :::|||
Db  1030 GAAGTAGAT----------AAAAATGGGCTCAAGATGTGTGAGCCTTGTGGGGGACTATGT  1080

Qy   123 GlyLysSerCys 126
            ||||:::|||
Db  1081 CCCAAAGCCTGT 1092
```

Search completed: October 24, 2004, 05:53:46
Job time : 94 secs

OM protein - nucleic search, using frame_plus_p2n model

```
          GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:    October 24, 2004, 02:50:16 ; Search time 432 Seconds
                                       (without alignments)
                                       1603.990 Million cell updates/sec

Title:     US-10-613-105-2
Perfect score: 762
Sequence:  1 MKSSGLFPFLVLLALGTLAP.........RDLKCCMGMCGKSCVSPVKA 132

Scoring table: BLOSUM62
           Xgapop  10.0 , Xgapext  0.5
           Ygapop  10.0 , Ygapext  0.5
           Fgapop   6.0 , Fgapext  7.0
           Delop    6.0 , Delext   7.0

Searched:  4134886 seqs, 262710521 residues

Total number of hits satisfying chosen parameters:    8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 100 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US10613105/runat_22102004_170440_18681/app_query.fasta_1.327
-DB=N Geneseq_23sep04 -QFMT=fastap -SUFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=100 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=45
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10613105 @CGN_1_1_470 @runat_22102004_170440_18681 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  N Geneseq_23sep04:*
         1:  geneseqn1980s:*
         2:  geneseqn1990s:*
         3:  geneseqn2000s:*
         4:  geneseqn2001as:*
         5:  geneseqn2001bs:*
         6:  geneseqn2002as:*
         7:  geneseqn2002bs:*
         8:  geneseqn2003as:*
         9:  geneseqn2003bs:*
        10:  geneseqn2003cs:*
        11:  geneseqn2003ds:*
        12:  geneseqn2004s:*
```

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	Query Match %	Length	DB	ID	Description
1	762	100.0	399	5	ABA83108	Aba83108 Secretory
2	762	100.0	422	10	ADE95609	Ade95609 Human NOV
3	762	100.0	551	10	ADF79666	Adf79666 Leukaemia
4	762	100.0	565	1	AAN70930	Aan70930 Sequence
5	762	100.0	565	4	AAF32043	Aaf32043 Human ant
6	762	100.0	594	6	ABS76518	Abs76518 cDNA enco
7	762	100.0	594	6	ABL67548	Abl67548 Thyroid c
8	762	100.0	594	6	ABL66679	Abl66679 Lung canc
9	762	100.0	594	6	AAS17499	Aas17499 Human CDN
10	762	100.0	594	6	ABT11084	Abt11084 Human bre
11	762	100.0	594	6	ABZ34724	Abz34724 Coding se
12	762	100.0	594	8	ACC50295	Acc50295 Breast ca
13	762	100.0	594	10	ADB75550	Adb75550 Prostate
14	762	100.0	594	10	ADE95611	Ade95611 Human NOV
15	762	100.0	594	10	ADF81611	Adf81611 Leukaemia
16	762	100.0	594	10	ACA56673	Aca56673 Human sig
17	762	100.0	594	11	ADK61171	Adk61171 Ovarian c
18	762	100.0	594	11	ADP65640	Adp65640 Human mRN
19	762	100.0	594	8	ADI56469	Adi56469 Human pol
20	762	100.0	594	10	ACC50296	Acc50296 Breast ca
21	762	100.0	598	11	ADB80589	Adb80589 Ovarian c
22	762	100.0	598	11	ADP65076	Adp65076 Human sec
23	762	100.0	598	12	ADL26771	Adl26771 Human SLP
24	762	100.0	598	12	ADN03925	Adn03925 Antipsori
25	762	100.0	599	2	AAZ77546	Aaz77546 Human ova
26	762	100.0	636	5	AAH57451	Aah57451 Human lun
27	762	100.0	1084	12	ADL62375	Adl62375 Human ova
28	756	99.2	594	12	ADL57146	Adl57146 Human NOV
29	755	99.1	594	12	ADL57144	Adl57144 Human NOV
30	754	99.0	594	12	ADL57140	Adl57140 Human NOV
31	750	98.4	594	12	ADL57142	Adl57142 Human NOV
32	750	97.8	567	10	ADL57148	Adl57148 Human NOV
33	745	94.9	594	12	ADK11941	Adk11941 Breast ca
34	723	94.6	594	9	ADL57150	Adl57150 Human NOV
35	720.5	92.3	605	2	AAL60894	Aal60894 Human sec
c 36	703	84.6	528	2	ADK61172	Adk61172 Ovarian c
37	645	84.6	460	4	AAX16271	Aax16271 DNA seque
38	645	84.6	460	4	AAC97622	Aac97622 DNA seque
39	645	83.8	460	6	AAI67582	Aai67582 DNA seque
40	638.5	83.3	1525	10	ABK88025	Abk88025 DNA seque
41	635	83.3	321	1	ABK88016	Abk88016 DNA encod
42	635	83.3	321	2	ADE95613	Ade95613 DNA encod
43	635	83.3	324	2	AAN60463	Aan60463 Serine pr
44	635	83.3	324	2	AAN60464	Aan60464 Synthetic
45	635	83.3	324	3	AAX16236	Aax16236 Serine pr
46	635	83.3	324	3	AAX16194	Aax16194 Serine pr
47	635	83.3	324	3	AAC97579	Aac97579 DNA encod
48	635	83.3	324	3	AAC97527	Aac97527 DNA encod
49	635	83.3	324	3	AAC97528	Aac97528 Secretory
50	635	83.3	324	4	AAI67488	Aai67488 Recombina
51	635	83.3	324	3	AAI67489	Aai67489 Recombina
52	635	83.3	436	3	AAC97624	Aac97624 DNA encod
53	635	83.3	436	6	AAI67583	Aai67583 DNA seque
54	635	83.3	1525	12	ABK88022	Abk88022 DNA seque
55	631	82.8	420	4	AAX16272	Aax16272 DNA seque
56	624	81.9	324	10	ADL57138	Adl57138 Human NOV
57	597	78.3	314	2	AAH57243	Aah57243 Human lun
58	597	78.3	314	12	ADK11942	Adk11942 Breast ca
59	596	78.2	321	1	AAC97529	Aac97529 DNA encod
60	557.5	73.2	1072	12	ADK70406	Adk70406 Respirato
61	547	71.8	2274	1	AAN60465	Aan60465 Synthetic
62	547	71.8	2274	2	AAC97530	Aac97530 Human gen
63	547	71.8	2274	4	AAI67491	Aai67491 Recombina
64	545	71.5	498	1	AAN70929	Aan70929 Sequence
65	539	70.7	920	1	AAN60468	Aan60468 ompA-secr
66	535	70.2	872	1	AAN60469	Aan60469 ompA-tc-m
c 67	463.5	60.8	396	6	ABQ73677	Abq73677 Murine SL
68	463.5	60.8	396	6	ABQ73674	Abq73674 Murine SL
69	463.5	60.8	409	8	ABZ34649	Abz34649 Coding se
70	463.5	60.8	691	2	AAX18514	Aax18514 Mouse IMC
71	458	60.1	1114	2	AAX18515	Aax18515 Mouse IMC
72	458	60.1	325	10	ACA56426	Aca56426 Human sig
73	458	60.1	325	12	ADI56222	Adi56222 Human pol
74	378	49.6	684	1	AAX18516	Aax18516 Mouse IMC
75	348	45.7	194	1	AAN90354	Aan90354 Sequence
76	346	45.4	180	3	AAC97526	Aac97526 DNA encod
77	346	45.4	183	4	AAX16273	Aax16273 Secretory
78	346	45.4	183	4	AAI67584	Aai67584 Secretory
79	339	44.5	180	2	AAQ45442	Aaq45442 CLPI. 3/2

```
80   339  44.5  AAZ88483  180   3  Aaz88483 Truncated
81   339  44.5  ACH04226  805   9  Ach04226 Human cDN
82   285  37.4  AAD45347  4090  6  Aad45347 Mouse ant
83   283.5 37.2 ADI70084  313   5  Adi70084 Human ova
84   283.5 37.2 ADI76414  313   3  Adi76414 Human ova
85   272  35.7  AAC97619  135   3  Aac97619 DNA inser
86   272  35.7  AAC97576  135   3  Aac97576 Insert B
87   272  35.7  AAI67538  135   4  Aai67538 Serine pr
88 c 264  34.6  ABL67811  411   6  Abl67811 Ovary can
89 c 264  34.6  ABL66677  411   6  Abl66677 Lung canc
90 c 264  34.6  ABL67461  411   6  Abl67461 Thyroid c
91 c 264  34.6  ABK64114  411   6  Abk64114 Human ben
92   257.5 33.8  ADL85843  248  12  Adl85843 DNA up-re
93   257.5 33.8  ADL85844  248  12  Adl85844 DNA up-re
94   243  31.9  ADL41650  188   5  Adl41650 Human ova
95   238  31.2  AAC97577  120   3  Aac97577 Insert C
96   238  31.2  AAC97620  120   4  Aac97620 DNA inser
97   238  31.2  AAI67539  120   4  Aai67539 Serine pr
98   224  29.4  AAA70066  292   3  Aaa70066 Human ova
99   224  29.4  ABN72960  292   6  Abn72960 Ovarian c
100  224  29.4  ADA08530  292   9  Ada08530 Human ova
```

                                    ALIGNMENTS

RESULT 1
ABA83108
ID ABA83108 standard; DNA; 399 BP.
XX
AC ABA83108;
XX
DT 08-FEB-2002 (first entry)
XX
DE Secretory leukocyte protease inhibitor ovarian tumour marker gene, #53.
XX
KW Ovarian tumour marker gene; human; overexpression; upregulation;
KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;
KW identification; serous cystadenoma; borderline serous tumour;
KW serous cystadenocarcinoma, mucinous cystadenocarcinoma;
KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma,
KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;
KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;
KW immune response pathway; cell proliferation regulation; protein folding;
KW membrane localised; secreted; therapeutic target; cytostatic;
KW gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WO200175177-A2.
XX
PD 11-OCT-2001.
XX
PF 03-APR-2001; 2001WO-US010947.
XX
PR 03-APR-2000; 2000US-0194336P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;
XX
DR WPI; 2001-626450/72.
DR P-PSDB; ABB50282.
XX
PT Detecting and identifying ovarian tumor, identifying increased risk for
PT developing ovarian cancer, and determining effectiveness of ovarian
PT cancer treatment, by measuring expression level of ovarian tumor marker
PT gene.
XX
PS Claim 23; Page 98; 140pp; English.
XX
CC The invention relates to methods for diagnosing and prognosing ovarian
CC tumours in an individual via the detection and measurement of the
CC expression of ovarian tumour marker genes (ABA83081-ABA83122, ABA83180,

CC ABA83182 and ABA83181 and ABA83183). The methods of the invention are useful for
CC detecting an ovarian tumour in a patient, for identifying an individual
CC at increased risk for developing ovarian cancer, in prognostic tests for
CC assessing the relative severity of ovarian cancer, in tests for
CC monitoring a patient in remission from ovarian cancer and in tests for
CC monitoring disease status in a patient being treated for ovarian cancer.
CC The methods can additionally be used to identify a particular tumour as
CC being an ovarian tumour (i.e., an epithelial ovarian tumour selected from
CC serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma,
CC mucinous cystadenoma, borderline mucinous tumour, mucinous
CC cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma,
CC clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner
CC tumour. The ovarian tumour marker genes of the invention were identified
CC using SAGE (serial analysis of gene expression) and were found to be
CC overexpressed in a broad variety of ovarian epithelial tumour cells
CC relative to normal ovarian epithelial cells. The marker genes are
CC implicated in immune response pathways, in the regulation of cell
CC proliferation and in protein folding, and many of these are membrane-
CC localised or secreted. In addition to their use as diagnostic and
CC prognostic markers, the ovarian tumour marker genes or their encoded
CC proteins may be used as therapeutic targets for the treatment and
CC prevention of ovarian cancer. Sequences ABA83081-ABA83122, ABA83180,
CC ABA83182 and ABA83184 represent the ovarian tumour marker genes of the
CC invention
XX
SQ Sequence 399 BP; 91 A; 100 C; 111 G; 97 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   1.3e-58   Length: 399
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 0

US-10-613-105-2 (1-132) x ABA83108 (1-399)

Qy   1 MetLysSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro   20
Db   1 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTGGCAACTCTGGCACCT   60

Qy 21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysSer   40
Db 61 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 120

Qy 41 AlaGlnCysLeuArgThrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
Db 121 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 180

Qy 61 LysArgCysCysProAspThrCysGlyGlyIleLysCysLeuAspProValAspThrProAsn 80
Db 181 AAGAGATGTTGTCCGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 240

Qy 81 ProThrArgArgLysProGlyLysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
Db 241 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTGATGCTTAAC 300

Qy 101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysMetGly 120
Db 301 CCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 360

Qy 121 MetCysGlyLysSerCysValSerProValLysAla 132
Db 361 ATGTGTGGGAAATCCTGCGTTTCCCTGTGAAAGCT 396

RESULT 2
ADE95609
ID ADE95609 standard; cDNA; 422 BP.
XX
AC ADE95609;
XX
DT 12-FEB-2004 (first entry)

```
XX
DE   Human NOVX27a protein cDNA sequence.
XX
KW   NOVX protein; biochemical stimulation; physiological stimulation;
KW   cardiant; antiarteriosclerotic; hypotensive; cytostatic; anorectic;
KW   antirheumatic; antiarthritic; antidiabetic; nephrotropic; dermatological;
KW   immunosuppressive; anti-HIV; antiinflammatory; neuroprotective;
KW   nootropic; antipsoriatic; antiparkinsonian; antiasthmatic; neuroleptic;
KW   antidepressant; antiallergic gynaecological; gene therapy; vaccine;
KW   NOVX-associated disorder; cardiomyopathy; atherosclerosis; hypertension;
KW   cancer; obesity; rheumatoid arthritis; diabetes; glomerulonephritis;
KW   psoriasis; skin disorder; AIDS; inflammation; multiple sclerosis;
KW   Alzheimer's disease; Parkinson's disease; asthma; schizophrenia;
KW   depression; allergy; fertility disorder; gene; ss; NOVX27a.
XX
OS   Homo sapiens.
XX
PN   WO200050245-A2.
XX
PD   19-JUN-2003.
XX
PF   03-DEC-2002; 2002WO-US038594.
XX
PR   05-DEC-2001; 2001US-0336600P.
PR   07-DEC-2001; 2001US-0338285P.
PR   12-DEC-2001; 2001US-0341346P.
PR   17-DEC-2001; 2001US-0341477P.
PR   17-DEC-2001; 2001US-0341540P.
PR   20-DEC-2001; 2001US-0342592P.
PR   27-DEC-2001; 2001US-0342297P.
PR   31-DEC-2001; 2001US-0344903P.
PR   17-APR-2002; 2002US-0373288P.
PR   15-MAY-2002; 2002US-0380981P.
PR   17-MAY-2002; 2002US-0381495P.
PR   28-MAY-2002; 2002US-0383534P.
PR   28-MAY-2002; 2002US-0387744P.
PR   29-MAY-2002; 2002US-0383829P.
PR   29-MAY-2002; 2002US-0384024P.
PR   07-AUG-2002; 2002US-0401788P.
PR   26-AUG-2002; 2002US-0406353P.
PR   31-OCT-2002; 2002US-00401788.
PR   02-DEC-2002; 2002US-00406353.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Alsobrook JP,  Anderson DW,  Boldog FL,  Burgess CE,  Chillakuru RA;
PI   Edinger SR,  Gerlach VL,  Gorman L,  Gould-Rothberg BB,  Guo X;
PI   Jeffers ME,  Ji W,  Li L,  Malyankar UM,  Miller CE,  Murphey R;
PI   Patturajan M,  Peyman JA,  Rastelli L,  Rieger DK,  Shenoy SG;
PI   Smithson G,  Starling G,  Taupier RJ,  Voss EZ,  Zhong H,  Zhong M;
XX
DR   WPI; 2003-513974/48.
DR   P-PSDB; ADE95610.
XX
PT   New NOVX polypeptides and nucleic acids, useful for preventing or
PT   treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT   atherosclerosis or diabetes, and in chromosome mapping, tissue typing or
PT   pharmacogenomics.
XX
PS   Claim 20; SEQ ID NO 141; 211pp; English.
XX
CC   This invention relates to novel NOVX proteins, and the DNA sequence which
CC   encode them, having properties related to stimulation of biochemical or
CC   physiological responses in a cell, a tissue, an organ or an organism.
CC   Compounds which modulate the proteins of the invention may have cardiant,
CC   antiarteriosclerotic, hypotensive, cytostatic, anorectic, antirheumatic,
CC   antiarthritic, antidiabetic, nephrotropic, dermatological,
CC   immunosuppressive, anti-HIV, antiinflammatory, neuroprotective,
CC   nootropic, antipsoriatic, antiparkinsonian, antiasthmatic, neuroleptic,
CC   antidepressant, antiallergic or gynaecological activities. The DNA
CC   sequences of the invention may be useful for gene therapy whilst the
CC   protein sequences may allow the development of a vaccine. The protein is
CC   useful in the manufacture of a medicament for treating a syndrome
```

```
CC   associated with a human disease. The invention may be useful in
CC   diagnosing, treating or preventing NOVX-associated disorders, for example
CC   cardiomyopathy, atherosclerosis, hypertension, cancer, obesity,
CC   rheumatoid arthritis, diabetes, glomerulonephritis, psoriasis, skin
CC   disorders, AIDS, inflammation, multiple sclerosis, Alzheimer's disease,
CC   Parkinson's disease, asthma, schizophrenia, depression, allergies or
CC   fertility disorders. The nucleic acids may further be used as
CC   hybridisation probes, in chromosome mapping, tissue typing, preventive
CC   medicine, and pharmacogenomics. The present sequence is the cDNA sequence
CC   which encodes the human NOVX27a protein of the invention.
XX
SQ   Sequence 422 BP; 94 A; 111 C; 119 G; 98 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:                                    422
Score:                          1.39e-58      Length:               422
Percent Similarity:             762.00        Matches:              132
Best Local Similarity:          100.00%       Conservative:         0
Query Match:                    100.00%       Mismatches:           0
DB:                             100.00%       Indels:               0
                                10            Gaps:                 0

US-10-613-105-2 (1-132) x ADB95609 (1-422)

Qy       1 MetLysSerSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro  20
          | | | | | | | | | | | | | | | | | | | |
Db      15 ATGAAGTCCAGCGGCCTTTTCCCCTTCCTGGTGCTGGCTCTGGGAACTCTGGCACCT  74

Qy      21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysSer  40
          | | | | | | | | | | | | | | | | | | | |
Db      75 TGGGCTGTGGAAGGCTCTGGAAAGTCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT  134

Qy      41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
          | | | | | | | | | | | | | | | | | | | |
Db     135 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGAAG  194

Qy      61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
          | | | | | | | | | | | | | | | | | | | |
Db     195 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC  254

Qy      81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyValGlnLeuMetLeuAsn  100
          | | | | | | | | | | | | | | | | | | | |
Db     255 CCAACAAGGAGGAAGCCTGGGAAGTGCCAGTGACTTATGGCCAATGTTGATGCTTAAC  314

Qy     101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly  120
          | | | | | | | | | | | | | | | | | | | |
Db     315 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC  374

Qy     121 MetCysGlyLysSerCysValSerProValLysAla  132
          | | | | | | | | | | | | |
Db     375 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  410
```

```
RESULT 3
ADF79666
ID   ADF79666 standard; DNA; 551 BP.
XX
AC   ADF79666;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Leukaemia-related DNA sequence #222.
XX
KW   Cytostatic; Gene therapy; leukaemia; ss.
XX
OS   Unidentified.
XX
PN   WO200303943-A2.
XX
PD   15-MAY-2003.
XX
PF   04-NOV-2002; 2002WO-EP012303.
XX
PR   05-NOV-2001; 2001EP-00126244.
PR   30-APR-2002; 2002EP-00009758.
```

XX
PA   (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA   (UVLU-) UNIV LUDWIG MAXIMILIANS.
PA   (HAFE/) HAFERLACH T.
PA   (SCHO/) SCHOCH C.
PA   (KERN/) KERN W.
XX
PI   Haferlach T, Schoch C, Kern W, Kohlmann A, Schnittger S, Dugas M;
PI   Eils R, Brors B, Mergenthaler S;
XX
DR   WPI; 2003-505037/47.
XX
PT   Determining the subtype of leukemia cells and whether a patient sample
PT   contains leukemia cells or other cells, useful for treating leukemia,
PT   comprises determining the expression profile of a group of markers in a
PT   patient sample.
XX
PS   Disclosure; SEQ ID NO 222; 2938pp; English.
XX
CC   The present invention relates to a method (M1) for determining the
CC   subtype of leukaemia cells and whether a patient sample contains
CC   leukaemia cells. The method comprises determining the expression profile
CC   of a group of markers in a patient sample. The method is useful for
CC   determining the presence of leukaemia cells, its types or subtypes, and
CC   for the preparation of a medicament for treating leukaemia.
XX
SQ   Sequence 551 BP; 117 A; 146 C; 150 G; 138 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   1.91e-58   Length: 551
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 10   Indels: 0
DB: Gaps: 0

US-10-613-105-2 (1-132) x ADF79666 (1-551)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   14 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTTGCCCTGGGAACTCTGGCCACCT 73

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyLysValCysProProLysLysSer 40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 133

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 134 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 193

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 194 AAGAGATGTTGTCTGACACCTGTGGCATCAAATGCCTAGATCCTGTTGACACCCCAAAC 253

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 254 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 313

Qy 101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysProCysLysMetGly 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 314 CCCCCCAATTTCTGTGAGATGGATGGCCAAGCGTGACTTGAAGTGTTGCATGGGC 373

Qy 121 MetCysGlyLysLysSerCysValSerProValLysAla 132
         ||||||||||||||||||||||||||||||||||||||||
Db 374 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 409

RESULT 4
AAN70930
ID   AAN70930 standard; cDNA; 565 BP.
XX
AC   AAN70930;
XX
DT   25-MAR-2003 (revised)

DT   21-MAY-1991 (first entry)
XX
DE   Sequence encoding protein with the biological activity of HUSI (human
DE   seminal plasma inhibitor) type I inhibitors encoded on pRH 34.
XX
KW   Bronchitis therapy; cervix inflammation; hyperfibrinolysis; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   59..457
FT   /*tag= a
XX
PN   DE360057l-A.
XX
PD   06-AUG-1987.
XX
PF   10-JAN-1986;   86DE-03600571.
XX
PR   10-JAN-1986;   86DE-03600571.
XX
PA   (CHEF) GRUENENTHAL GMBH.
XX
PI   Heinzel R, Appelhans H, Gassen HG, Machleidt W;
XX
DR   WPI; 1987-222038/32.
DR   P-PSDB; AAP70584.
XX
PT   New protein with human seminal plasma inhibitor activity - and new DNA
PT   coding sequences, vectors and transformed organisms, useful e.g. for
PT   treating bronchitis or inflammation.
XX
PS   Claim 5; Fig 5; 28pp; German.
XX
CC   HUSI-type I inhibitors are useful for treating chronic bronchitis,
CC   chronic cervical inflammation; inflammation associated with excessive
CC   mucus prodn.; post-operative bleeding caused by hyperfibrinolysis, and
CC   shock. They are esp. suitable for administration as spray or inhalation.
CC   The expression control sequence on the DNA fragment in AAN70928 is bound
CC   to the 5' end of HUSI cDNA in an expression vector. (Updated on 25-MAR-
CC   2003 to correct PA field.)
XX
SQ   Sequence 565 BP; 106 A; 162 C; 171 G; 126 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   1.96e-58   Length: 565
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 0

US-10-613-105-2 (1-132) x AAN70930 (1-565)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   59 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTTGCCCTGGGAACTCTGGCCACCT 118

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyLysValCysProProLysSer 40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 119 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 178

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 179 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 238

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 239 AAGAGATGTTGTCTGACACCTGTGGCATCAAATGCCTAGATCCTGTTGACACCCCAAAC 298

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 299 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 358

```
Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  359 CCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 418

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
        ||||||||||||||||||||||||||||||||||||
Db  419 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 454

RESULT 5
AAF32043
ID  AAF32043 standard; cDNA; 565 BP.
XX
AC  AAF32043;
XX
DT  10-APR-2001  (first entry)
XX
DE  Human antileukoprotease coding sequence.
XX
KW  Human; cytostatic; gynecological; elastase; cathepsin G; gene therapy;
KW  endometriosis; antileukoprotease; ALP; secretory leukoprotease inhibitor;
KW  SLPI; alpha1-proteinase inhibitor; alpha1-PI; ss.
XX
OS  Homo sapiens.
XX
PN  WO200101998-A2.
XX
PD  11-JAN-2001.
XX
PF  29-JUN-2000; 2000WO-IB000990.
XX
PR  01-JUL-1999; 99US-0142157P.
PR  27-JUN-2000; 2000US-0605134.
XX
PA  (REPR-) REPROGEN INC.
XX
PI  Neilson L, Li Z;
XX
DR  WPI; 2001-138062/14.
DR  P-PSDB; AAB66864.
XX
PT  Treating and diagnosing endometriosis or promoting embryo implantation
PT  comprises determining or modulating antileukoprotease activity in a
PT  patient.
XX
PS  Disclosure; Fig 1; 28pp; English.
XX
CC  The present invention relates to a method for treating endometriosis in a
CC  subject, comprising administering to the subject an amount of
CC  antileukoprotease (ALP) effective to inhibit the activity of elastase or
CC  cathepsin G, thereby inhibiting ectopic implantation of endometrial
CC  fragments. The present sequence is the coding sequence for human ALP. ALP
CC  is also known as secretory leukoprotease inhibitor (SLPI) and alpha1-
CC  proteinase inhibitor (alpha1-PI). ALP plays an important role in the
CC  defence of epithelial surfaces against proteolytic damage
XX
SQ  Sequence 565 BP; 106 A; 162 C; 171 G; 126 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:
Score:                  1.96e-58
                        762.00
Percent Similarity:     100.00%
Best Local Similarity:  100.00%
Query Match:            100.00%
DB:                     4

US-10-613-105-2 (1-132) x AAF32043 (1-565)

Qy    1 MetLysSerGlyLysSerGlyLysLeuPheProPheLeuValLeuAlaLeuAlaGlyThrLeuAlaPro 20
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   59 ATGAAGTCCAGCGGCGCCTCTTCCCCTTCCTGGTCGTCGCCCTAGGAACTCTGGCACCT 118

Qy   21 TrpAlaValGluGlyGlyLysSerGlyLysSerPheLysAlaGlyValValCysProLysLysSer 40
```

```
Db  119 TGGGCTGTGGAAGGCTGGAAAGTCCTTCAAAGCTGGAGAGTCGTCCTCCTAAGAAATCT 178
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   41 AlaGlnCysLeuArgTyrLysProGluCysGlnSerAspTrpProGlnCysProGlyLys 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  179 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTCCGACTGGCCAGTCCAGGGAAG 238

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysGlyLeuAspProValAspThrProAsn 80
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  239 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCTGGACTCCGATCCTGTTGACACCCCAAAC 298

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  299 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACCTATGGCCAATGTTTGATGCTTAAC 358

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  359 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACCTGACTGGCATGGGC 418

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
        ||||||||||||||||||||||||||||||||||||
Db  419 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 454

RESULT 6
ABS76518
ID  ABS76518 standard; cDNA; 594 BP.
XX
AC  ABS76518;
XX
DT  11-DEC-2002  (first entry)
XX
DE  cDNA encoding human ovarian cancer marker M185.
XX
KW  Human; ovarian cancer; marker; cancer; familial history; brain disorder;
KW  central nervous system disorder; bacterial meningitis; viral meningitis;
KW  Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
KW  brain herniation; inflammation; encephalitis; testicular disorder;
KW  nontuberculous granulomatous orchitis; connective tissue disorder;
KW  heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
KW  histological type; carcinogenic; ovarian cancer marker; gene; ss.
XX
OS  Homo sapiens.
XX
PN  WO200271928-A2.
XX
PD  19-SEP-2002.
XX
PF  14-MAR-2002; 2002WO-US007826.
XX
PR  14-MAR-2001; 2001US-0276025P.
PR  14-MAR-2001; 2001US-0276026P.
PR  10-AUG-2001; 2001US-0311732P.
PR  19-SEP-2001; 2001US-0323580P.
PR  26-SEP-2001; 2001US-0324967P.
PR  26-SEP-2001; 2001US-0325102P.
PR  26-SEP-2001; 2001US-0325149P.
XX
PA  (MILL-) MILLENNIUM PHARM INC.
XX
PI  Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
PI  Meyers RE, Morrisey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
PI  Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;
XX
DR  WPI; 2002-723277/78.
DR  P-PSDB; ABG96419.
XX
PT  Assessing whether a patient is afflicted with ovarian cancer, useful in
PT  assessing the stage or progression of the disease, comprises comparing
PT  the expression level of a cancer marker in a sample from a patient and
PT  from a non cancer patient.
XX
PS  Disclosure; Page 410-411; 481pp; English.
XX
```

CC   The present invention relates to a new method for assessing whether a
CC   patient is afflicted with ovarian cancer. The method involves comparing
CC   the expression level of a marker in a patient sample and the normal level
CC   of expression of the marker in a control non-ovarian cancer sample, where
CC   the marker is selected from 363 cancer markers described in the
CC   specification. The method of the invention is useful in diagnosing or
CC   characterising cancer, in detecting the presence of cancer as early as
CC   possible, and the recurrence of ovarian cancer. The method may also be of
CC   particular use with patients having an enhanced risk of developing
CC   ovarian cancer (e.g. patients having a familial history of ovarian
CC   cancer). The cancer markers may be used in the management and treatment
CC   of e.g. brain and central nervous system disorders (e.g. bacterial and
CC   viral meningitis, Alzheimer's disease or Parkinson's disease), brain
CC   disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
CC   inflammations (e.g. bacterial or viral meningitis or encephalitis),
CC   testicular disorders (e.g. nontuberculous granulomatous orchitis),
CC   connective tissue disorders, or heart disorders (e.g. ischaemic heart
CC   disease or atherosclerosis). The compositions and methods may also be
CC   used in assessing the histological type of neoplasm associated with
CC   ovarian cancer, monitoring the progression of ovarian cancer, determining
CC   whether ovarian cancer has metastasized or is likely to metastasize,
CC   selecting a composition for inhibiting ovarian cancer, assessing the
CC   ovarian carcinogenic potential of a compound, or inhibiting ovarian
CC   cancer or at risk of developing ovarian cancer. The present nucleic acid
CC   sequence encodes one of the ovarian cancer markers described in the
CC   invention
XX
SQ   Sequence 594 BP; 132 A; 156 C; 155 G; 151 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:
Score: 2.08e-58
                         762.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 6

Length: 594
Matches: 132
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-613-105-2 (1-132) x ABS76518 (1-594)

```
Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  19 ATGAAGTCCAGCGGCCCTTCCCCTTCCCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCT  78

Qy  21 TrpAlaValGluGlyPheGlyLysSerPheLysAlaGlyValCysProProLysLysSer  40
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  79 TGGGCTGTGGAAGGCTTCGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 138

Qy  41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy  61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy  81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACGTATGGCCAATGTTTGATGCTTAAC 318

Qy 101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy 121 MetCysGlyLysSerCysValSerProValLysAla 132
         ||||||||||||||||||||||||||||||||||||
Db 379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414
```

RESULT 7
ABL67548
ID   ABL67548 standard; DNA; 594 BP.
XX
AC   ABL67548;
XX

DT   15-MAY-2002 (first entry)
XX
DE   Thyroid cancer related gene sequence SEQ ID NO:5885.
XX
KW   Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW   stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW   cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO20194629-A2.
XX
PD   13-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US010838.
XX
PR   05-JUN-2000; 2000US-0209473P.
PR   05-JUN-2000; 2000US-0209531P.
PR   18-SEP-2000; 2000US-0231133P.
PR   18-SEP-2000; 2000US-0236617P.
PR   20-SEP-2000; 2000US-0234009P.
PR   20-SEP-2000; 2000US-0234034P.
PR   20-SEP-2000; 2000US-0234052P.
PR   22-SEP-2000; 2000US-0234509P.
PR   22-SEP-2000; 2000US-0234567P.
PR   25-SEP-2000; 2000US-0234923P.
PR   25-SEP-2000; 2000US-0234924P.
PR   25-SEP-2000; 2000US-0235077P.
PR   25-SEP-2000; 2000US-0235082P.
PR   25-SEP-2000; 2000US-0235114P.
PR   25-SEP-2000; 2000US-0235134P.
PR   25-SEP-2000; 2000US-0235280P.
PR   26-SEP-2000; 2000US-0235637P.
PR   26-SEP-2000; 2000US-0235638P.
PR   27-SEP-2000; 2000US-0235711P.
PR   27-SEP-2000; 2000US-0235720P.
PR   27-SEP-2000; 2000US-0235840P.
PR   27-SEP-2000; 2000US-0235863P.
PR   28-SEP-2000; 2000US-0236028P.
PR   28-SEP-2000; 2000US-0236032P.
PR   28-SEP-2000; 2000US-0236033P.
PR   28-SEP-2000; 2000US-0236034P.
PR   28-SEP-2000; 2000US-0236109P.
PR   28-SEP-2000; 2000US-0236111P.
PR   29-SEP-2000; 2000US-0236842P.
PR   29-SEP-2000; 2000US-0236891P.
PR   02-OCT-2000; 2000US-0237172P.
PR   02-OCT-2000; 2000US-0237173P.
PR   02-OCT-2000; 2000US-0237278P.
PR   02-OCT-2000; 2000US-0237294P.
PR   02-OCT-2000; 2000US-0237295P.
PR   02-OCT-2000; 2000US-0237316P.
PR   02-OCT-2000; 2000US-0237425P.
PR   03-OCT-2000; 2000US-0237598P.
PR   03-OCT-2000; 2000US-0237604P.
PR   03-OCT-2000; 2000US-0237606P.
PR   03-OCT-2000; 2000US-0237608P.
PR   01-NOV-2000; 2000US-0244867P.
PR   01-NOV-2000; 2000US-0245084P.
XX
PA   (AVAL-) AVALON PHARM.
XX
PI   Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI   Soppet DR, Weaver Z;
XX
DR   WPI; 2002-188264/24.
XX
PT   Screening for anti-neoplastic agent involves exposing cells to a chemical
PT   agent to be tested for anti-neoplastic activity, and determining a change
PT   in expression of a gene of a signature gene set.
XX
PS   Claim 1; SEQ ID NO 5885; 44pp; English.
XX

CC   The present invention describes a method (M1) for screening for an anti-
CC   neoplastic agent. The method involves exposing cells to a chemical agent
CC   to be tested for anti-neoplastic activity, determining a change in
CC   expression of at least one gene (I) of a signature gene set, where (I)
CC   comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC   to ABL70110), or is at least 95% identical to (S), where a change in
CC   expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC   activity and can be used in gene therapy. M1 can be used for screening an
CC   anti-neoplastic agent, and can be used for producing a product which is
CC   the data collected with respect to the anti-neoplastic agent as a result
CC   of M1, and the data is sufficient to convey the chemical structure and/or
CC   properties of the agent. M1 can be used in the treatment of cancer such
CC   as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC   prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC   cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC   cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC   tumour
XX
SQ   Sequence 594 BP; 132 A; 156 C; 155 G; 151 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   2.08e-58   Length: 594
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 6

US-10-613-105-2 (1-132) x ABL67548 (1-594)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCT 78

Qy   21 TrpAlaValGluGlyGlyLysSerGlyLysSerPheLysLysLysValCysProProLysLysSer 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 138

Qy   41 AlaGlnCysLeuArgTyrLArgLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyLysMetLeuAsn 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   259 CCAACAAGGAGGAAGCCTGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy   101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   319 CCCCCAAATTTCTGTGAGATGGATGGCCAGTGCAAGCCGTGACTTGAAGTGTTGCATGGGC 378

Qy   121 MetCysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||||||
Db   379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414

RESULT 8
ABL66679
ID   ABL66679 standard; DNA; 594 BP.
XX
AC   ABL66679;
XX
DT   15-MAY-2002 (first entry)
XX
DE   Lung cancer related gene sequence SEQ ID NO:5016.
XX
KW   Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW   stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW   cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW   gene; ds.
XX

OS   Homo sapiens.
XX
PN   WO200194629-A2.
XX
PD   13-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US010838.
XX
PR   05-JUN-2000; 2000US-0209473P.
PR   05-JUN-2000; 2000US-0209531P.
PR   18-SEP-2000; 2000US-0233133P.
PR   18-SEP-2000; 2000US-0233617P.
PR   20-SEP-2000; 2000US-0234009P.
PR   20-SEP-2000; 2000US-0234034P.
PR   20-SEP-2000; 2000US-0234052P.
PR   22-SEP-2000; 2000US-0234509P.
PR   22-SEP-2000; 2000US-0234567P.
PR   25-SEP-2000; 2000US-0234923P.
PR   25-SEP-2000; 2000US-0234924P.
PR   25-SEP-2000; 2000US-0235077P.
PR   25-SEP-2000; 2000US-0235082P.
PR   25-SEP-2000; 2000US-0235134P.
PR   25-SEP-2000; 2000US-0235280P.
PR   26-SEP-2000; 2000US-0235637P.
PR   26-SEP-2000; 2000US-0235638P.
PR   27-SEP-2000; 2000US-0235711P.
PR   27-SEP-2000; 2000US-0235720P.
PR   27-SEP-2000; 2000US-0235840P.
PR   27-SEP-2000; 2000US-0235863P.
PR   28-SEP-2000; 2000US-0236028P.
PR   28-SEP-2000; 2000US-0236032P.
PR   28-SEP-2000; 2000US-0236033P.
PR   28-SEP-2000; 2000US-0236034P.
PR   28-SEP-2000; 2000US-0236109P.
PR   28-SEP-2000; 2000US-0236111P.
PR   29-SEP-2000; 2000US-0236842P.
PR   29-SEP-2000; 2000US-0236891P.
PR   02-OCT-2000; 2000US-0237172P.
PR   02-OCT-2000; 2000US-0237173P.
PR   02-OCT-2000; 2000US-0237278P.
PR   02-OCT-2000; 2000US-0237294P.
PR   02-OCT-2000; 2000US-0237295P.
PR   02-OCT-2000; 2000US-0237316P.
PR   02-OCT-2000; 2000US-0237425P.
PR   03-OCT-2000; 2000US-0237598P.
PR   03-OCT-2000; 2000US-0237604P.
PR   03-OCT-2000; 2000US-0237606P.
PR   03-OCT-2000; 2000US-0237608P.
PR   01-NOV-2000; 2000US-0244867P.
PR   01-NOV-2000; 2000US-0245084P.
XX
PA   (AVAL-) AVALON PHARM.
XX
PI   Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI   Soppet DR, Weaver Z;
XX
DR   WPI; 2002-188264/24.
XX
PT   Screening for anti-neoplastic agent involves exposing cells to a chemical
PT   agent to be tested for anti-neoplastic activity, and determining a change
PT   in expression of a gene of a signature gene set.
XX
PS   Claim 1; SEQ ID NO 5016; 44pp; English.
XX
CC   The present invention describes a method (M1) for screening for an anti-
CC   neoplastic agent. The method involves exposing cells to a chemical agent
CC   to be tested for anti-neoplastic activity, determining a change in
CC   expression of at least one gene (I) of a signature gene set, where (I)
CC   comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC   to ABL70110), or is at least 95% identical to (S), where a change in
CC   expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC   activity and can be used in gene therapy. M1 can be used for screening an
CC   anti-neoplastic agent, and can be used for producing a product which is

Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 6

US-10-613-105-2 (1-132) x ABT11084 (1-594)

```
Qy      1 MetLysSerSerGlyLeuProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro    20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     19 ATGAAGTCCAGCGGCCTCTTCCCTTCCCTTCTGGTTGCTGCTGCCCTGGAACCTGGCACCT    78

Qy     21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProLysLysSer    40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     79 TGGGCTGTGGAAGGCTCTGGAAAGTCTTCAAAGCTGGAGTCTGTCCCTAAGAAATCT    138

Qy     41 AlaGlnCysLeuArgThrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG    198

Qy     61 LysArgCysCysProAspThrCysGlyIleLysCysGlyLeuAspProValAspThrProAsn    80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    199 AAGAGATGTTGTCCTGACACACTTGTGGCATCAAATGCCTGGACTCCTGTTGACACCCCAAAC    258

Qy     81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyLnCysLeuMetLeuAsn    100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC    318

Qy    101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysLysMetGly    120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC    378

Qy    121 MetCysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||||||||||||
Db    379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414
```

RESULT 11
ABZ34724
ID ABZ34724 standard; cDNA; 594 BP.
XX
AC ABZ34724;
XX
DT 04-FEB-2003 (first entry)
XX
DE Coding sequence SEQ ID 82, upregulated in osteogenesis.
XX
KW Osteopathic; osteogenesis modulator; gene therapy; osteogenesis;
KW osteoporosis; bone disease; downregulator; human; ss.
XX
OS Homo sapiens.
XX
PN WO200281745-A2.
XX
PD 17-OCT-2002.
XX
PF 05-APR-2002; 2002WO-IB002211.
PR 05-APR-2001; 2001US-0281400P.
XX
PA (AVET) AVENTIS PHARMA SA.
XX
PI Garcia T, Roman Roman S, Baron R, Call K, Theilhaber J;
PI Connolly T, Jackson A, Bushnell SE, Rawadi G;
XX
DR WPI; 2003-058567/05.
XX
PT Novel isolated nucleic acid upregulated/downregulated in osteogenesis,
PT useful for bone disease therapy in subject.
XX
PS Claim 26; Page 116; 237pp; English.
XX
CC The present invention relates to novel nucleotide sequences, which are
CC differentially expressed in models of osteogenesis upon being put in
CC contact with a stimulator of osteogenesis. The present sequence is one
```
    199 AAGAGATGTTGTCCTGACACACTTGTGGCATCAAATGCCTGGACTCCTGTTGACACCCCAAAC 258
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnThrTyrGlyGlnCysLeuMetLeuAsn 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318
```

```
Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysLysMetGly 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
        ||||||||||||||||||||||||||||||||||||
Db  379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414
```

RESULT 10
ABT11084
ID ABT11084 standard; cDNA; 594 BP.
XX
AC ABT11084;
XX
DT 04-DEC-2002 (first entry)
XX
DE Human breast cancer associated coding sequence SEQ ID NO: 1218.
XX
KW Human; breast specific gene; breast cancer; differential expression;
KW cytostatic; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200259271-A2.
XX
PD 01-AUG-2002.
XX
PF 25-JAN-2002; 2002WO-US002176.
XX
PR 25-JAN-2001; 2001US-0263757P.
PR 25-APR-2001; 2001US-0286090P.
PR 23-MAY-2001; 2001US-0292517P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Orr MS, Nation M, Diggans JC, Zeng W;
XX
DR WPI; 2002-674803/72.
XX
PT Diagnosing breast cancer in a patient comprises detecting the level of
PT gene expression in cell or tissue samples, where a differential gene
PT expression is indicative of breast cancer.
XX
PS Claim 1; SEQ ID NO 1218; 260pp + Sequence Listing; English.
XX
CC The present invention relates to methods of diagnosing breast cancer in a
CC patient, which comprise detecting the level of expression in a tissue
CC sample of two or more genes selected from those shown in ABT09867-
CC ABT11112, where a differential expression of the genes indicates breast
CC cancer. The methods are useful in diagnosing, treating, detecting the
CC progression, and in monitoring treatment of breast cancer in patients.
CC The methods are also useful as a screening tool for agents that modulate
CC the onset or progression of breast cancer. The breast cancer genes may be
CC used as diagnostic markers for the prediction or identification of the
CC malignant state of breast tissue, for confirming the type and progression
CC of cancer, and for drug screening and assays. The present sequence is a
CC coding sequence of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub.published_pct_sequences
XX
SQ Sequence 594 BP; 132 A; 156 C; 155 G; 151 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   2.08e-58
Score: 762.00

Length: 594
Matches: 132

CC such sequence. This sequence can be used for diagnosing osteoporosis/bone
CC disease in a patient, promoting osteogenesis and/or preventing
CC osteoporosis/bone disease. The present sequence encodes a secreted
CC protein
XX
SQ Sequence 594 BP; 132 A; 156 C; 155 G; 151 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   2.08e-58   Length: 594
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 8   Gaps: 0

US-10-613-105-2 (1-132) x ABZ34724 (1-594)

```
Qy    1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 ATGAAGTCCAGCGGCCGCCTCTTCCCCTTCCTGTGCTTGCCCTGGAACTCTGGCACCT   78

Qy   21 TrpAlaValGluGlyLysSerPheLysSerPheLysAlaGlyLysValCysProLysSer  40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGTCGGAGTCGTCCTCCTAAGAAATCT  138

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTGACTGGCAGTGTCCAGGGAAG  198

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC  258

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLysLeuMetLeuAsn  100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC  318

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGTGCAAGGCTGACTGACTTGAAGTGTTGCATGGGC  378

Qy  121 MetCysGlyLysLysSerCysValSerProValLysAla  132
          |||||||||||||||||||||||||||||||
Db  379 ATGTGTGGAAAATCCTGCGTTTCCCCTGTGAAAGCT  414
```

RESULT 12
ACC50295
ID   ACC50295 standard; cDNA; 594 BP.
XX
AC   ACC50295;
XX
DT   12-JUN-2003 (first entry)
XX
DE   Breast cancer associated cDNA sequence SEQ ID NO:429.
XX
KW   Human; breast cancer; cytostatic; gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003004989-A2.
XX
PD   16-JAN-2003.
XX
PF   21-JUN-2002; 2002WO-US019669.
XX
PR   21-JUN-2001; 2001US-0299887P.
PR   27-JUN-2001; 2001US-0301572P.
PR   18-JUL-2001; 2001US-0305501P.
PR   25-SEP-2001; 2001US-0325002P.
PR   05-MAR-2002; 2002US-0362585P.
PR   14-MAY-2002; 2002US-0380391P.
XX
PA   (MILL-) MILLENIUM PHARM INC.

XX
PI   Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S;
PI   Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;
PI   Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;
XX
DR   WPI; 2003-210381/20.
DR   P-PSDB; ABR47595.
XX
PT   Breast cancer diagnosis or treatment by comparing the level of expression
PT   of a marker in a patient sample with that in the control non-breast
PT   cancer sample.
XX
PS   Claim 1; SEQ ID NO 429; 128pp; English.
XX
CC   The present invention describes a method for assessing whether a patient
CC   is afflicted with breast cancer. The method comprises comparing the level
CC   of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
CC   ABR47386 to ABR47632) in a patient sample and the normal level of
CC   expression of the marker in a control non-breast cancer sample, where a
CC   significant increase in the level of expression of the marker in the
CC   patient sample and the normal level is an indication that the patient is
CC   afflicted with breast cancer. The breast cancer associated sequences from
CC   the present invention have cytostatic activities and can be used in gene
CC   therapy. The method is useful for diagnosing and treating breast cancer.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 594 BP; 132 A; 156 C; 155 G; 151 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   2.08e-58   Length: 594
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 8   Gaps: 0

US-10-613-105-2 (1-132) x ACC50295 (1-594)

```
Qy    1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 ATGAAGTCCAGCGGCCCTTCCCCTTCCTGTGCGTGCTTGCCCTGGAACTCTGGCACCT   78

Qy   21 TrpAlaValGluGlyLysSerPheLysSerPheLysAlaGlyLysValCysProLysSer  40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGTCGGAGTCGTCCTCCTAAGAAATCT  138

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTGACTGGCAGTGTCCAGGGAAG  198

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC  258

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLysLeuMetLeuAsn  100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC  318

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGTGCAAGGCTGACTGACTTGAAGTGTTGCATGGGC  378

Qy  121 MetCysGlyLysLysSerCysValSerProValLysAla  132
          |||||||||||||||||||||||||||||||
Db  379 ATGTGTGGAAAATCCTGCGTTTCCCCTGTGAAAGCT  414
```

RESULT 13
ADB75550
ID   ADB75550 standard; cDNA; 594 BP.
XX
AC   ADB75550;

```
Qy    61  LysArgCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn    80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   199  AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC   258

Qy    81  ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn   100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   259  CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC   318

Qy   101  ProProAsnPheCysGluMetCysGlyGlnCysLysLysArgAspLeuLysCysMetGly   120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   319  CCCCCCAATTTCTGTGAGATGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC   378

Qy   121  MetCysGlyLysSerCysValSerProValLysAla   132
             ||||||||||||||||||||||||||||||||||||
Db   379  ATGTGTGGGAAAATCCTGCGTTTCCCCTGTGAAAGCT   414
```

RESULT 14
ADE95611
ID   ADE95611 standard; cDNA; 594 BP.
XX
AC   ADE95611;
XX
DT   12-FEB-2004 (first entry)
XX
DE   Human NOVX27b protein cDNA sequence.
XX
KW   NOVX protein; biochemical stimulation; physiological stimulation;
KW   cardiant; antiarteriosclerotic; hypotensive; cytostatic; anorectic;
KW   antirheumatic; antiarthritic; antidiabetic; nephrotropic; dermatological;
KW   immunosuppressive; anti-HIV; antiinflammatory; neuroprotective;
KW   nootropic; antipsoriatic; antiparkinsonian; antiasthmatic; neuroleptic;
KW   antidepressant; antiallergic; gynaecological; gene therapy; vaccine;
KW   NOVX-associated disorder; cardiomyopathy; atherosclerosis; hypertension;
KW   cancer; obesity; rheumatoid arthritis; diabetes; glomerulonephritis;
KW   psoriasis; skin disorder; AIDS; inflammation; multiple sclerosis;
KW   Alzheimer's disease; Parkinson's disease; asthma; schizophrenia;
KW   depression; allergy; fertility disorder; gene; SS; NOVX27b.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   variation   replace(38,C)
FT   /*tag= a
FT   variation   replace(53,C)
FT   /standard_name= "Single nucleotide polymorphism"
FT   /*tag= b
FT   variation   replace(169,C)
FT   /standard_name= "Single nucleotide polymorphism"
FT   /*tag= c
FT   variation   replace(173,G)
FT   /standard_name= "Single nucleotide polymorphism"
FT   /*tag= d
FT   variation   replace(188,A)
FT   /standard_name= "Single nucleotide polymorphism"
FT   /*tag= e
FT   variation   replace(193,A)
FT   /standard_name= "Single nucleotide polymorphism"
FT   /*tag= f
FT   /standard_name= "Single nucleotide polymorphism"
XX
PN   WO200305024 5-A2.
XX
PD   19-JUN-2003.
XX
PF   03-DEC-2002; 2002WO-US038594.
XX
PR   05-DEC-2001; 2001US-0336600P.
PR   07-DEC-2001; 2001US-0338285P.
PR   12-DEC-2001; 2001US-0341346P.
PR   17-DEC-2001; 2001US-0341477P.
PR   17-DEC-2001; 2001US-0341540P.
PR   20-DEC-2001; 2001US-0342592P.
PR   27-DEC-2001; 2001US-0344297P.
```

```
XX
DT   04-DEC-2003  (first entry)
XX
DE   Prostate cancer marker cDNA.
XX
KW   Prostate; cancer; cytostatic; gene therapy; marker; SS.
XX
OS   Homo sapiens.
XX
PN   WO200309814-A2.
XX
PD   06-FEB-2003.
XX
PF   25-JUL-2002; 2002WO-US023913.
XX
PR   25-JUL-2001; 2001US-0307982P.
PR   22-AUG-2001; 2001US-0314356P.
PR   25-SEP-2001; 2001US-0325020P.
PR   12-DEC-2001; 2001US-0341746P.
PR   05-MAR-2002; 2002US-0362158P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Schlegel R, Monahan JE, Endege WO, Gannavarapu M,  Gorbatcheva B;
PI   Hoersh S, Kamatkar S, Wonsey AM, Glatt K, Zhao X,  Anderson D;
XX
DR   WPI; 2003-248033/24.
XX
PT   New nucleic acid molecule, useful for diagnosing or treating prostate
PT   cancer.
XX
PS   Disclosure; SEQ ID NO 374; 99pp; English.
XX
CC   The invention relates to newly discovered cancer markers associated with
CC   the cancerous state of prostate cells. Also disclosed is a method of
CC   assessing whether a patient is afflicted with prostate cancer. The method
CC   of the invention involves assessing whether a patient is afflicted with
CC   prostate cancer by comparing the level of expression of a marker in a
CC   patient sample and the normal level of expression of the marker in a
CC   control non-prostate cancer sample, where a significant increase in the
CC   level of expression of the marker in the patient sample and the normal
CC   level indicates that the patient is afflicted with prostate cancer.
CC   Nucleic acids of the invention are useful for diagnosing or treating
CC   prostate cancer, and may be useful in gene therapy. Sequences given in
CC   ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 594 BP; 132 A; 156 C; 155 G; 151 T; 0 U; 0 Other;
```

Alignment Scores:
```
Pred. No.:          2.08e-58       Length:            594
Score:              762.00         Matches:           132
Percent Similarity: 100.00%        Conservative:        0
Best Local Similarity: 100.00%     Mismatches:          0
Query Match:        100.00%        Indels:              0
Db:                 10             Gaps:                0
```

US-10-613-105-2 (1-132) x ADB75550 (1-594)

```
Qy     1  MetLysSerSerGlyLeuPheProPheLeuValLeuValAlaLeuGlyThrLeuAlaPro    20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    19  ATGAAGTCCAGCGACGCGGCCTTCCCCCTTCCCGTGCTCGTCGCCCTGGGAACTCTGGCACCT    78

Qy    21  TrpAlaValGluGlySerGlyLysSerPheLysSerPheLysAlaGlyValCysProPro LysLysSer    40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    79  TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT   138

Qy    41  AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys    60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   139  GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTCAGTGACTGGCAGTGTCCAGGGAAG   198
```

```
PR   31-DEC-2001; 2001US-0344903P.
PR   17-APR-2002; 2002US-0373288P.
PR   15-MAY-2002; 2002US-0380981P.
PR   17-MAY-2002; 2002US-0381409P.
PR   28-MAY-2002; 2002US-0383534P.
PR   28-MAY-2002; 2002US-0383744P.
PR   29-MAY-2002; 2002US-0383829P.
PR   29-MAY-2002; 2002US-0384024P.
PR   07-AUG-2002; 2002US-0401788P.
PR   26-AUG-2002; 2002US-0406533P.
PR   31-OCT-2002; 2002US-00401788.
PR   02-DEC-2002; 2002US-00406353.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Chillakuru RA;
PI   Edinger SR, Gerlach VL, Gorman L, Gould-Rothberg BE, Guo X;
PI   Jeffers ME, Ji W, Li L, Malyankar UM, Miller CE, Murphey R;
PI   Patturajan M, Peyman JA, Rastelli L, Rieger DK, Shenoy SG;
PI   Smithson G, Starling G, Taupier RJ, Voss EZ, Zhong H, Zhong M;
XX
DR   WPI; 2003-513974/48.
DR   P-PSDB; ADE95612.
XX
PT   New NOVX polypeptides and nucleic acids, useful for preventing or
PT   treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT   atherosclerosis or diabetes, and in chromosome mapping, tissue typing or
PT   pharmacogenomics.
XX
PS   Claim 20; SEQ ID NO 143; 21pp; English.
XX
CC   This invention relates to novel NOVX proteins, and the DNA sequence which
CC   encode them, having properties related to stimulation of biochemical or
CC   physiological responses in a cell, a tissue, an organ or an organism.
CC   Compounds which modulate the proteins of the invention may have cardiant,
CC   antiarteriosclerotic, hypotensive, cytostatic, anorectic, antirheumatic,
CC   antiarthritic, antidiabetic, nephrotropic, dermatological,
CC   immunosuppressive, anti-HIV, antiinflammatory, neuroprotective,
CC   nootropic, antipsoriatic, antiparkinsonian, antiasthmatic, neuroleptic,
CC   antidepresant, antiallergic or gynaecological activities. The DNA
CC   sequences of the invention may be useful for gene therapy whilst the
CC   protein sequences may allow the development of a vaccine. The protein is
CC   useful in the manufacture of a medicament for treating a syndrome
CC   associated with a human disease. The invention may be useful in
CC   diagnosing, treating or preventing NOVX-associated disorders, for example
CC   cardiomyopathy, atherosclerosis, hypertension, cancer, obesity,
CC   rheumatoid arthritis, diabetes, glomerulonephritis, psoriasis, skin
CC   disorders, AIDS, inflammation, multiple sclerosis, Alzheimer's disease,
CC   Parkinson's disease, asthma, schizophrenia, depression, allergies or
CC   fertility disorders. The nucleic acids may further be used as
CC   hybridisation probes, in chromosome mapping, tissue typing, preventive
CC   medicine, and pharmacogenomics. The present sequence is the cDNA sequence
CC   which encodes the human NOVX27b protein of the invention.
XX
SQ   Sequence 594 BP; 132 A; 156 C; 155 G; 151 T; 0 U; 0 Other;

     Alignment Scores:
     Pred. No.:            2.08e-58      Length:             594
     Score:                762.00        Matches:            132
     Percent Similarity:   100.00%       Conservative:       0
     Best Local Similarity: 100.00%      Mismatches:         0
     Query Match:          100.00%       Indels:             0
     DB:                                 Gaps:               10

US-10-613-105-2 (1-132) x ADE95611 (1-594)

Qy      1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     19 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCT  78

Qy     21 TrpAlaValGluGlySerGlyLysSerPheLysSerAlaGlyValCysProProLysSer  40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGTCGGAGTCTGTCCTCCTAAGAAATCT 138
```

```
Qy     41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy     61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy     81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    259 CCAACAAGGAGGAAGCCTGGAAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy    101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy    121 MetCysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||||||||||||
Db    379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414

RESULT 15
ADF81611
ID   ADF81611 standard; DNA; 594 BP.
XX
AC   ADF81611;
XX
DT   26-FEB-2004   (first entry)
XX
DE   Leukaemia-related DNA sequence #2167.
XX
KW   Cytostatic; Gene therapy; leukaemia; ss.
XX
OS   Unidentified.
XX
PN   WO200303943-A2.
XX
PD   15-MAY-2003.
XX
PF   04-NOV-2002; 2002WO-EP012303.
XX
PR   05-NOV-2001; 2001EP-00126244.
PR   30-APR-2002; 2002EP-00009758.
XX
PA   (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA   (UYLU-) UNIV LUDWIG MAXIMILIANS.
PA   (HAFB/) HAFERLACH T.
PA   (SCHO/) SCHOCH C.
PA   (KERN/) KERN W.
XX
PI   Haferlach T, Schoch C, Kern W, Kohlmann A, Schnitiger S, Dugas M;
PI   Eils R, Brors B, Mergenthaler S;
XX
DR   WPI; 2003-505037/47.
XX
PT   Determining the subtype of leukemia cells and whether a patient sample
PT   contains leukemia cells or other cells, useful for treating leukemia,
PT   comprises determining the expression profile of a group of markers in a
PT   patient sample.
XX
PS   Disclosure; SEQ ID NO 2167; 2938pp; English.
XX
CC   The present invention relates to a method (M1) for determining the
CC   subtype of leukaemia cells and whether a patient sample contains
CC   leukaemia cells. The method comprises determining the expression profile
CC   of a group of markers in a patient sample. The method is useful for
CC   determining the presence of leukaemia cells, its types or subtypes, and
CC   for the preparation of a medicament for treating leukaemia.
XX
SQ   Sequence 594 BP; 132 A; 156 C; 155 G; 151 T; 0 U; 0 Other;

     Alignment Scores:
     Pred. No.:            2.08e-58      Length:             594
```

```
Score:                    762.00      Matches:            132
Percent Similarity:       100.00%     Conservative:         0
Best Local Similarity:    100.00%     Mismatches:           0
Query Match:                          Indels:               0
DB:                                   Gaps:

US-10-613-105-2 (1-132) x ADF81611 (1-594)

Qy      1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     19 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTGGCTCTGGGAACTCTGGCACCT  78

Qy     21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer  40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT  138

Qy     41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerSerAspTrpGlnCysProGlyLys  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG  198

Qy     61 LysArgCysProAspThrCysGlyIleLeuCysLeuAspProValAspProValAspThrProAsn  80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    199 AAGAGATGTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC  258

Qy     81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyCysLeuMetLeuAsn  100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGCGCCAATGTTTGATGCTTAAC  318

Qy    101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysMetGly  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    319 CCCCCCAATTTCTGTGAGATGGATGGCCAAGCGTGCAAGCGTGACTTGAAGTGTTGCATGGGC  378

Qy    121 MetCysGlyLysSerCysValSerProValLysAla  132
          ||||||||||||||||||||||||||||||||||||
Db    379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  414

RESULT 16
ACA56673
ID   ACA56673 standard; cDNA; 594 BP.
XX
AC   ACA56673;
XX
DT   06-JUN-2003 (first entry)
XX
DE   Human signalling pathway polynucleotide probe SEQ ID NO 1271.
XX
KW   Human; probe; ss; array element; Parkinson's disease;
KW   signalling pathway population; cancer; adenocarcinoma; leukaemia;
KW   immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
XX
OS   Homo sapiens.
XX
PN   US6500938-B1.
XX
PD   31-DEC-2002.
XX
PF   30-JAN-1998;  98US-00016434.
XX
PR   30-JAN-1998;  98US-00016434.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Au-Young J, Seilhamer JJ;
XX
DR   WPI; 2003-352189/33.
XX
PT   Combination of polynucleotide probes, useful as array elements in a
PT   microarray for monitoring the expression of a number of target
PT   polynucleotides.
XX
PS   Claim 1; SEQ ID NO 1271; 65pp; English.
XX
CC   The invention relates to a combination which, comprises a number of
```

```
Score:                    762.00      Matches:            132
Percent Similarity:       100.00%     Conservative:         0
Best Local Similarity:    100.00%     Mismatches:           0
Query Match:                          Indels:               0
DB:                                   Gaps:

CC   polynucleotide probes comprising a sequence selected from one of the 1490
CC   sequences mentioned in the specification. The combination is useful as an
CC   array element in a microarray for monitoring the expression of a number
CC   of target polynucleotides. The microarray is particularly useful in the
CC   diagnosis and treatment of cancer and immunopathy and neuropathology.
CC   The microarray is useful in diagnostics and treatment regimens, drug
CC   discovery and development, toxicological and carcinogenicity studies,
CC   forensics and pharmacogenomics. The microarray is also useful for
CC   monitoring progression of diseases and for developing sophisticated
CC   profiles for the effects of currently available therapeutic drugs. The
CC   combination is also useful for purifying a subpopulation of mRNAs, cDNAs
CC   and genomic fragments and in research and diagnostic applications. The
CC   array can detect changes in expression in a large number of genes coding
CC   for different signaling pathway populations which can be used to diagnose
CC   various diseases including cancer e.g. adenocarcinoma and leukaemia,
CC   immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
CC   and Parkinson's disease. The present sequence represents a polynucleotide
CC   probe of the invention. Note: The sequence data for this patent did not
CC   form part of the printed specification but was obtained in electronic
CC   format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=0650093B1
XX
SQ   Sequence 594 BP; 132 A; 156 C; 155 G; 151 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                             Length:              594
Score:                    2.08e-58     Matches:             132
Percent Similarity:       762.00       Conservative:          0
Best Local Similarity:    100.00%      Mismatches:            0
Query Match:              100.00%      Indels:                0
DB:                       10           Gaps:                  0

US-10-613-105-2 (1-132) x ACA56673 (1-594)

Qy      1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     19 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTGGCTCTGGGAACTCTGGCACCT  78

Qy     21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer  40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT  138

Qy     41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerSerAspTrpGlnCysProGlyLys  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG  198

Qy     61 LysArgCysProAspThrCysGlyIleLeuCysLeuAspProValAspProValAspThrProAsn  80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    199 AAGAGATGTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC  258

Qy     81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyCysLeuMetLeuAsn  100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGCGCCAATGTTTGATGCTTAAC  318

Qy    101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysMetGly  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    319 CCCCCCAATTTCTGTGAGATGGATGGCCAAGCGTGCAAGCGTGACTTGAAGTGTTGCATGGGC  378

Qy    121 MetCysGlyLysSerCysValSerProValLysAla  132
          ||||||||||||||||||||||||||||||||||||
Db    379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  414

RESULT 17
ADK61171
ID   ADK61171 standard; DNA; 594 BP.
XX
AC   ADK61171;
XX
DT   06-MAY-2004 (first entry)
XX
DE   Ovarian cancer-related DNA #326 with altered ovarian cancer expression.
XX
KW   ds; gene; ovarian tumor; BRCA-1-like; BRCA-2-like; non-BRCA-like;
```

KW gene expression; primer; cancer.
OS Homo sapiens.
XX
PN WO200306054-A2.
XX
PD 21-AUG-2003.
XX
PF 13-FEB-2003; 2003WO-US004688.
XX
PR 13-FEB-2002; 2002US-0357031P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (SLOK) SLOAN KETTERING INST CANCER RES.
XX
PI Jazaeri AA, Boyd J, Liu ET;
XX
DR WPI; 2003-689589/65.
XX
PT Classifying an ovarian tumor as a BRCA-1-like or BRCA-2-
PT like tumor by determining a pattern of expression in the ovarian tumor of
PT several markers.
XX
PS Disclosure; SEQ ID NO 341; 137pp; English.
XX
CC The invention relates to a method of classifying an ovarian tumor as a
CC BRCA-1-like or BRCA-2-like or non-BRCA-like tumor by: (1) determining a
CC pattern of expression in the ovarian tumor of several markers given in
CC the specification; and (2) comparing a similarity of the pattern of
CC expression of the markers in the ovarian tumor to a pattern of expression
CC of the markers in a comparison tissue of a known BRCA-1-like or BRCA-2-
CC like or non-BRCA-like tumor. The method is useful for classifying an
CC ovarian tumor as a BRCA-1-like or BRCA-2-like or non-BRCA-like tumor.
CC This sequence corresponds to an ovarian cancer-related gene having an
CC altered pattern of expression in ovarian cancer. (Note: The sequence data
CC for this patent did not form part of the printed specification but was
CC obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences).
XX
SQ Sequence 594 BP; 132 A; 156 C; 155 G; 151 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   2.08e-58
Score: 762.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 10

	Length:	594
Matches:	132
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-10-613-105-2 (1-132) x ADK61171 (1-594)

```
Qy      1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     19 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTTGCGCTTGGAACCCTGGCACCT  78

Qy     21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer  40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 138

Qy     41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy     61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTTGATCCTGTTGACACCCCAAAC 258

Qy     81 ProThrArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    259 CCAACAAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy    101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Db   319 CCCCCAATTCTGTGAGATGGATGGCCAGTGCAAGCGTGCAAGGTGCTTGAAGTGTTGCATGGGC 378

Qy   121 MetCysGlyLysSerValSerProValLysAla 132
          |||||||||||||||||||||||||||||||||
Db   379 ATGTGTGGGAAATCCGTTTCCCCTGTGAAAGCT 414

RESULT 18
ADP65640
ID ADP65640 standard; DNA; 594 BP.
XX
AC ADP65640;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human mRNA for antileukoprotease (ALP) from cervix uterus DNA.
XX
KW autoimmune disease; arthritis; gene expression analysis;
KW rheumatoid arthritis; collagen-induced; immunosuppressive; antirheumatic;
KW antiarthritic; osteopathic; antigout; antiinflammatory; dermatological;
KW immunomodulatory; lupus; ankylosing spondylitis; Fibrositis;
KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;
KW immune; ds; human.
XX
OS Homo sapiens.
XX
PN WO200307827-A1.
XX
PD 04-SEP-2003.
XX
PF 31-OCT-2002; 2002WO-US035433.
XX
PR 31-OCT-2001; 2001US-0336220P.
XX
PA (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
XX
PI Hirsch R, Thorton SL;
XX
DR WPI; 2003-712740/67.
DR GENBANK; X04470.
XX
PT Diagnosing and analyzing autoimmune disease using gene expression
PT profiles and microarray technology, useful for diagnosing and treating
PT rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and
PT gout.
XX
PS Disclosure; Page; 56pp; English.
XX
CC The invention relates to a novel method for diagnosing and analysing
CC autoimmune disease or arthritides. The method comprises obtaining a
CC patient sample containing mRNA, analysing gene expression using the mRNA
CC that results in a gene expression signature of the mRNA, and using that
CC gene expression signature to diagnose or analyse the autoimmune disease
CC or arthrides in the patient, where the gene expression is of at least 60% of
CC the genes correlates with that of the gene signature. The invention
CC further comprises: a treatment of rheumatoid arthritis; identification of
CC genes for targeting in the treatment of rheumatoid arthritis in a mammal
CC other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an
CC array or gene chip, specific for rheumatoid arthritis; diagnosis or
CC analyses of autoimmune disease or rheumatoid arthritis; screening the
CC efficacy of a candidate drug in vitro for the treatment of collagen-
CC induced arthritis; and reducing the symptoms associated with collagen-
CC induced arthritis. The compositions of the invention have the following
CC activities: immunosuppressive, antirheumatic, antiarthritic, osteopathic,
CC antigout, antiinflammatory, dermatological, and immunomodulatory. The
CC methods and compositions of the present invention are useful for
CC diagnosing and treating autoimmune disease or arthritides, such as
CC rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,
CC fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an
CC immune disease caused by an infectious agent. This polynucleotide
CC represents a DNA sequence relating to the genes used in the analysis and
CC treatment of autoimmune diseases or arthritides. Note: This sequence is
CC not shown in the specification. It has been supplied in an electronic
CC format from WIPO.

DR　WPI: 2004-090520/09.
XX
PT　New composition comprising polynucleotide probes, useful as array
PT　elements in a microarray for monitoring the expression of target
PT　polynucleotides or purifying a subpopulation of mRNAs, cDNA, or genomic
PT　fragments.
XX
PS　Claim 6; SEQ ID NO 1271; 73pp; English.
XX
CC　The invention relates to a composition of polynucleotide probes
CC　comprising first polynucleotide probes comprising at least a portion of a
CC　gene encoding a receptor-like polypeptide, second polynucleotide probes
CC　comprising at least a portion of a gene encoding a transducing
CC　polypeptide and third polynucleotide probes comprising at least a portion
CC　of a gene encoding an effector-like polypeptide. The probes of the
CC　composition are useful as array elements in a microarray for monitoring
CC　the expression of target polynucleotides. The microarray is useful in the
CC　diagnosis and treatment of cancer, an immunopathology or a
CC　neuropathology. It can also be used for drug discovery and development,
CC　toxicological and carcinogenicity studies, forensics or pharmacogenomics.
CC　Microarrays can also be used for monitoring the progression of diseases
CC　that may be associated with the altered expression of signalling pathway
CC　polypeptides. The composition can also be used to purify a subpopulation
CC　of mRNAs, cDNAs, or genomic fragments in a sample. The expression profile
CC　is also useful for the diagnosis and treatment of cancer, e.g. cancers of
CC　the adrenal gland, bladder, bone, bone marrow, brain, breast or cervix,
CC　an immunopathology, e.g. AIDS, diabetes, pancreatitis, osteoporosis or
CC　ulcerative colitis, or a neuropathology, e.g. dementia, amnesia,
CC　epilepsy, Alzheimer's disease or depression. This sequence represents a
CC　human polynucleotide probe of the invention. Note: The sequence data for
CC　this patent did not form part of the printed specification but was
CC　obtained in electronic format directly from USPTO at
CC　seqdata.uspto.gov/sequence.html.
XX
SQ　Sequence 594 BP; 132 A; 156 C; 155 G; 151 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:　　　　　　　　　　　Length:　　　　　594
Score:　　　　2.08e-58　　　　　Matches:　　　　132
Percent Similarity:　100.00%　　Conservative:　　0
Best Local Similarity:　100.00%　Mismatches:　　　0
Query Match:　　　　　　　　　Indels:　　　　　0
DB:　　　　　　　　　　　　　Gaps:　　　　　　0

US-10-613-105-2 (1-132) x ADI56469 (1-594)

Qy　　1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
Db　 19 ATGAAGTCCAGCGGCCTCTTCCCTTCCCTTCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCT 78

Qy　 21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer 40
Db　 79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 138

Qy　 41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnSerAspTrpProGlyLys 60
Db　139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCCAGTGTCCAGGGAAG 198

Qy　 61 LysArgCysCysProAspThrCysGlyIleLeuLysCysLeuAspAspProValAspThrProAsn 80
Db　199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy　 81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
Db　259 CCAACAAGGAGGAAGCCTGGAAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy　101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysMetGly 120
Db　319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy　121 MetCysGlyLysSerCysValSerProValLysAla 132
Db　379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414

RESULT 19
ADI56469
ID　ADI56469 standard; DNA; 594 BP.
XX
AC　ADI56469;
XX
DT　22-APR-2004 (first entry)
XX
DE　Human polynucleotide probe #1271.
XX
KW　Human; probe; ss; receptor-like polypeptide; transducing polypeptide;
KW　effector-like polypeptide; cancer; immunopathology; neuropathology;
KW　drug development; toxicology; carcinogenicity;
KW　signalling pathway polypeptide; adrenal gland; bladder; bone;
KW　bone marrow; brain; breast; cervix; tumour; immunopathology; AIDS;
KW　diabetes; pancreatitis; osteoporosis; ulcerative colitis; neuropathology;
KW　dementia; amnesia; epilepsy; Alzheimer's disease; depression.
XX
OS　Homo sapiens.
XX
PN　US2004010136-A1.
XX
PD　15-JAN-2004.
XX
PF　26-NOV-2002; 2002US-00305720.
XX
PR　30-JAN-1998; 98US-00016434.
XX
PA　(INCY-) INCYTE GENOMICS INC.
XX
PI　Au-Young J, Seilhamer JJ;
XX

RESULT 20
ACC50296
ID   ACC50296 standard; cDNA; 598 BP.
XX
AC   ACC50296;
XX
DT   12-JUN-2003 (first entry)
XX
DE   Breast cancer associated cDNA sequence SEQ ID NO:431.
XX
KW   Human; breast cancer; cytostatic; gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003004989-A2.
XX
PD   16-JAN-2003.
XX
PF   21-JUN-2002; 2002WO-US019669.
XX
PR   21-JUN-2001; 2001US-0299887P.
PR   27-JUN-2001; 2001US-0301572P.
PR   18-JUL-2001; 2001US-0306501P.
PR   25-SEP-2001; 2001US-0325002P.
PR   05-MAR-2002; 2002US-0362585P.
PR   14-MAY-2002; 2002US-0380391P.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S;
PI   Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;
PI   Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;
XX
DR   WPI; 2003-210381/20.
XX
PT   Breast cancer diagnosis or treatment by comparing the level of expression
PT   of a marker in a patient sample with that in the control non-breast
PT   cancer sample.
XX
PS   Claim 1; SEQ ID NO 431; 128pp; English.
XX
CC   The present invention describes a method for assessing whether a patient
CC   is afflicted with breast cancer. The method comprises comparing the level
CC   of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
CC   ABR47386 to ABR47632) in a patient sample and the normal level of
CC   expression of the marker in a control non-breast cancer sample, where a
CC   significant increase in the level of expression of the marker in the
CC   patient sample and the normal level is an indication that the patient is
CC   afflicted with breast cancer. The breast cancer associated sequences from
CC   the present invention have cytostatic activities and can be used in gene
CC   therapy. The method is useful for diagnosing and treating breast cancer.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 598 BP; 134 A; 157 C; 156 G; 151 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:		Length:	598
Score:	2.1e-58	Matches:	762.00
		Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	100.00%	Gaps:	0
DB:			8

US-10-613-105-2 (1-132) x ACC50296 (1-598)

```
Qy    1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   23 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTGGCCCTGGGAACTCTGGCACCT 82

Qy   21 TrpAlaValGluGluSerGlyLysSerPheLysSerGlyValAlaGlyValCysProLysLysSer 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   83 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGTGGAGTCTGTCCTCCTAAGAAATCT 142

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  143 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTCTGACTGGCAGTGTCCAGGGAAG 202

Qy   61 LysArgCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  203 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTGACACCCCAAAC 262

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlnCysLeuMetLeuAsn 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  263 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 322

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysMetGly 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  323 CCCCCAATTTCTGTGAGATGATGGCCAGTGCAAGCGTGACTTGACTTGAGTGTTGCATGGGC 382

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||||||||||||
Db  383 ATGTGTGGAAATCCTGCGTTTCCCCTGTGAAAGCT 418
```

RESULT 21
ADB80589
ID   ADB80589 standard; DNA; 598 BP.
XX
AC   ADB80589;
XX
DT   04-DEC-2003 (first entry)
XX
DE   Ovarian cancer-associated transcript #78.
XX
KW   cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
KW   post-operative chemotherapy; radiation therapy; tumour prognosis;
KW   pre-cancerous lesion detection; ds; gene.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   23..421
FT   /*tag= a
XX
PN   WO2002102235-A2.
XX
PD   27-DEC-2002.
XX
PF   18-JUN-2002; 2002WO-US019297.
XX
PR   18-JUN-2001; 2001US-0299234P.
PR   27-AUG-2001; 2001US-0315287P.
PR   05-SEP-2001; 2001US-0317544P.
PR   13-NOV-2001; 2001US-0350666P.
PR   12-APR-2002; 2002US-0372246P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Mack DH, Gish KC;
XX
DR   WPI; 2003-167431/16.
DR   P-PSDB; ADB80590.
XX
PT   Detecting an ovarian cancer-associated transcript in a cell from a
PT   patient, comprises contacting a biological sample from the patient with a
PT   polynucleotide that hybridizes to an ovarian cancer gene.
XX
PS   Claim 10; Page 327; 332pp; English.
XX
CC   The invention relates to a method of detecting an ovarian cancer-
CC   associated transcript in a cell from a patient, by contacting a
CC   biological sample from the patient with a polynucleotide that selectively
CC   hybridizes to a sequence at least 80% identical to any one of 80
CC   nucleic acid sequences given in the specification. The method is useful

PF   31-OCT-2002; 2002WO-US035433.
XX
PR   31-OCT-2001; 2001US-0336220P.
XX
PA   (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
XX
PI   Hirsch R, Thorton SL;
XX
DR   WPI; 2003-712740/67.
DR   GENBANK; NM_003064.
XX
PT   Diagnosing and analyzing autoimmune disease using gene expression
PT   profiles and microarray technology, useful for diagnosing and treating
PT   rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and
PT   gout.
XX
PS   Disclosure; Page; 56pp; English.
XX
CC   The invention relates to a novel method for diagnosing and analysing
CC   autoimmune disease or arthritides. The method comprises obtaining a
CC   patient sample containing mRNA, analysing gene expression using the mRNA
CC   that results in a gene expression signature of the mRNA, and using that
CC   gene expression signature to diagnose or analyse the autoimmune disease
CC   or arthritides in the patient, where gene expression of at least 60% of
CC   the genes correlates with that of the gene signature. The invention
CC   further comprises: a treatment of rheumatoid arthritis; identification of
CC   genes for targeting in the treatment of rheumatoid arthritis in a mammal
CC   other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an
CC   array or gene chip, specific for rheumatoid arthritis; diagnosis or
CC   analyses of autoimmune disease or rheumatoid arthritis; screening the
CC   efficacy of a candidate drug in vitro for the treatment of collagen-
CC   induced arthritis; and reducing the symptoms associated with collagen-
CC   induced arthritis. The compositions of the invention have the following
CC   activities: immunosuppressive, antirheumatic, antiarthritic, osteopathic,
CC   antigout, antiinflammatory, dermatological, and immunomodulatory. The
CC   methods and compositions of the present invention are useful for
CC   diagnosing and treating autoimmune disease or arthritides, such as
CC   rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,
CC   fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an
CC   immune disease caused by an infectious agent. This polynucleotide
CC   represents a DNA sequence relating to the genes used in the analysis and
CC   treatment of autoimmune diseases or arthritides. Note: This sequence is
CC   not shown in the specification. It has been supplied in an electronic
CC   format from WIPO.
XX
SQ   Sequence 598 BP; 134 A; 157 C; 156 G; 151 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   2.1e-58   Length: 598
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: Gaps: 0

US-10-613-105-2 (1-132) x ADP65076 (1-598)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   23 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCT 82

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer 40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   83 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 142

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   143 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 202

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   203 AAGAGATGTTGCCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 262

CC   in diagnosing ovarian cancer and in identifying and using agents and/or
CC   targets that inhibit ovarian cancer. The nucleic acid molecule,
CC   polypeptide and the antibody may also be used in detecting ovarian
CC   cancers, monitoring and early detection of relapse following treatment,
CC   monitoring response to therapy, selecting patients for post-operative
CC   chemotherapy or radiation therapy, in selecting mode of therapy,
CC   determining tumour prognosis, early detection of pre-cancerous lesions,
CC   and as vaccines. This sequence corresponds to one of the nucleic acids
CC   used for the detection method of the invention.
XX
SQ   Sequence 598 BP; 134 A; 157 C; 156 G; 151 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   2.1e-58   Length: 598
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 10   Gaps: 0

US-10-613-105-2 (1-132) x ADB80589 (1-598)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   23 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCT 82

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer 40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   83 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 142

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   143 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 202

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   203 AAGAGATGTTGCCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 262

Qy   81 ProThrArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263 CCAACAAGGAGGAAGCCTGGAAGTGCCCAGTGACTTATGCCAATGTTTGATGCTTAAC 322

Qy   101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly 120
         ||||||||||||||||||||||||||||||||||||||||||||
Db   323 CCCCCAAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTGTGCATGGGC 382

Qy   121 MetCysGlyLysSerCysGlySerValSerProValLysAla 132
         |||||||||||||||||||||||||||||||||||
Db   383 ATGTGTGGGAAATCCTGCGGTTCCCCTGTGAAAGCT 418

RESULT 22
ADP65076
ID   ADP65076 standard; DNA; 598 BP.
XX
AC   ADP65076;
XX
DT   12-AUG-2004 (first entry)
XX
DE   Human secretory leukocyte protease inhibitor DNA sequence.
XX
KW   autoimmune disease; arthritide; gene expression analysis;
KW   rheumatoid arthritis; collagen-induced; immunosuppressive; antirheumatic;
KW   antiarthritic; osteopathic; antigout; antiinflammatory; dermatological;
KW   immunomodulatory; lupus; ankylosing spondylitis; fibrositis;
KW   fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;
KW   immune; ds; human.
XX
OS   Homo sapiens.
XX
PN   WO2003072827-A1.
XX
PD   04-SEP-2003.
XX

```
Qy    81  ProThrArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263  CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 322

Qy   101  ProProAsnPheCysGluMetAspGlyCysGlnCysLysArgAspLeuLysCysMetGly 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   323  CCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 382

Qy   121  MetCysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||||||||||
Db   383  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 418
```

RESULT 23
ADL26771
ID   ADL26771 standard; cDNA; 598 BP.
XX
AC   ADL26771;
XX
DT   03-JUN-2004 (first entry)
XX
DE   Human SLPI encoding cDNA SEQ ID NO:25.
XX
KW   ovarian cancer; ovarian cancer-associated transcript; cytostatic;
KW   gene therapy; human; secretory leukocyte protease inhibitor;
KW   chromosome 20; gene; 88.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   23..421
FT   /*tag= a
FT   /product= "SLPI"
XX
PN   WO2004022778-A1.
XX
PD   18-MAR-2004.
XX
PF   05-SEP-2003; 2003WO-AU001166.
PR   05-SEP-2002; 2002AU-00951346.
XX
PA   (GARV-) GARVAN INST MEDICAL RES.
XX
PI   Sutherland R, Henshall S, O'brien P;
XX
DR   WPI; 2004-315574/29.
DR   P-PSDB; ADL26772.
XX
PT   Use of genes and proteins for diagnosing ovarian cancer and/or a
PT   likelihood for survival or recurrence of the disease.
XX
PS   Claim 2; SEQ ID NO 25; 447pp; English.
XX
CC   The present invention describes a method for the use of genes and
CC   proteins for diagnosing ovarian cancer and/or a likelihood for survival
CC   or recurrence of the disease, where the expression of genes and proteins
CC   is up-regulated and down-regulated, associated with the occurrence or
CC   recurrence of a specific cancer sub-type. Also described: (1) detecting
CC   an ovarian cancer-associated transcript in a biological sample; (2)
CC   diagnosing an ovarian cancer in a human or animal subject being tested;
CC   (3) detecting an ovarian cancer-associated polypeptide in a biological
CC   sample; (4) monitoring the efficacy of a therapeutic treatment of ovarian
CC   cancer; (5) determining the likelihood of survival of a subject suffering
CC   from an ovarian cancer; and (6) an assay device for use in the diagnosis
CC   or prognosis of ovarian cancer comprising polynucleotides or antibodies
CC   immobilised to a solid phase, where each of the polynucleotides consists
CC   of a gene given in the specification and each of the antibodies binds to
CC   a polypeptide also given in the specification; and identifying a
CC   candidate compound for the treatment of ovarian cancer. An ovarian cancer
CC   -associated sequence has cytostatic activity, and can be used in gene
CC   therapy. An ovarian cancer-associated polynucleotide, vector, polypeptide
CC   or antibody can be used for the diagnosis or prognosis of ovarian cancer
CC   or for the preparation of a medicament for the treatment of ovarian

CC   cancer. The ovarian cancer that is diagnosed is an epithelial ovarian
CC   cancer selected from serous ovarian cancer, non-invasive ovarian cancer,
CC   mixed phenotype ovarian cancer, mucinous ovarian cancer, endometrial
CC   ovarian cancer, clear cell ovarian cancer, papillary serous ovarian
CC   cancer, Brenner cell or undifferentiated adenocarcinoma. The present
CC   sequence encodes human secretory leukocyte protease inhibitor (SLPI),
CC   which is located on chromosome 20 and is used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 598 BP; 134 A; 157 C; 156 G; 151 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   2.1e-58   Length: 598
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 12   Gaps: 0

US-10-613-105-2 (1-132) x ADL26771 (1-598)
```

```
Qy     1  MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    23  ATGAAGTCCAGCGGCCTTCCTTCCCTTCCTGGTGCTCCTGGCCCTTGCCCTGGAACTCTGGCACCT 82

Qy    21  TrpAlaValGluGlySerPheLysSerGlyValCysProLysLysSer 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    83  TGGGCTGTGGGAAGGCTCTGGAAAGTCCTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 142

Qy    41  AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   143  GCCCAGTGCCTTAGATCACAGAAACCTGAGTGCCAGTGACTGGCAGTGTCCAGGGAAG 202

Qy    61  LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   203  AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 262

Qy    81  ProThrArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263  CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 322

Qy   101  ProProAsnPheCysGluMetAspGlyCysGlnCysLysArgAspLeuLysCysMetGly 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   323  CCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 382

Qy   121  MetCysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||||||||||
Db   383  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 418
```

RESULT 24
ADN03925
ID   ADN03925 standard; cDNA; 598 BP.
XX
AC   ADN03925;
XX
DT   01-JUL-2004 (first entry)
XX
DE   Antipsoriatic cDNA sequence #162.
XX
KW   ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004028479-A2.
XX
PD   08-APR-2004.
XX
PF   25-SEP-2003; 2003WO-US030907.
PR   25-SEP-2002; 2002US-0414006P.
XX
PA   (GETH) GENENTECH INC.
XX

```
PI  Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI  Wu TD;
XX
DR  WPI; 2004-305105/28.
DR  P-PSDB; ADN03926.
XX
PT  New PRO nucleic acid or polypeptide, useful for preparing a
PT  pharmaceutical composition for diagnosing or treating psoriasis in a
PT  mammal.
XX
PS  Claim 1; SEQ ID NO 319; 3069pp; English.
XX
CC  The invention relates to novel polynucleotide and polypeptides for
CC  treating psoriasis or a sequence having at least 80% identity to the
CC  above sequences. The nucleic acid is useful for preparing a composition
CC  for diagnosing or treating psoriasis in a mammal. This sequence
CC  corresponds to one of the polynucleotides of the invention.
XX
SQ  Sequence 598 BP; 134 A; 157 C; 156 G; 151 T; 0 U; 0 Other;
XX

Alignment Scores:
Pred. No.:             2.1e-58         Length:         598
Score:                 762.00         Matches:        132
Percent Similarity:    100.00%        Conservative:   0
Best Local Similarity: 100.00%        Mismatches:     0
Query Match:           100.00%        Indels:         0
DB:                                   Gaps:           12

US-10-613-105-2 (1-132) x ADN03925 (1-598)

Qy    1 MetLysSerSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro  20
Db   23 ATGAAGTCCAGCGGCCCTTCCCCTTCCTGCGTGCTGGCTCTGGGAACTCTGGCACCT  82
Qy   21 TrpAlaValGluGlyLysSerGlyLysSerPheLysAlaGlyLysValCysProLysSer  40
Db   83 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCGTCCTCCTAAGAAATCT 142
Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
Db  143 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 202
Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
Db  203 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 262
Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyLysCysLeuMetLeuAsn 100
Db  263 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 322
Qy  101 ProProAsnPheCysGluMetAspGlyLysCysLysArgAspLeuLysCysCysMetGly 120
Db  323 CCCCCAAATTTCTGTGAGATGGATGGCAAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 382
Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
Db  383 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 418

RESULT 25
AA277546
ID  AA277546 standard; cDNA; 599 BP.
XX
AC  AA277546;
XX
DT  10-APR-2000 (first entry)
XX
DE  Human ovarian tumor cDNA library derived EST fragment 97.
XX
KW  Expressed sequence tag; EST; human; ovarian tumor; anticancer;
KW  gene therapy; treatment; ss.
XX
OS  Homo sapiens.
XX
```

```
PN  DE19817557-A1.
XX
PD  21-OCT-1999.
XX
PF  09-APR-1998;   98DE-01017557.
XX
PR  09-APR-1998;   98DE-01017557.
PA  (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI  Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX
DR  WPI; 1999-591920/51.
XX
PT  New nucleic acid sequences expressed in ovarian, and some other, cancer
PT  tissues, and derived polypeptides, for treatment of ovarian cancer and
PT  identification of therapeutic agents.
XX
PS  Claim 3; Page 221; 310pp; German.
XX
CC  This invention describes novel nucleic acid (cDNA) sequences (A) which
CC  have anticancer activity and are highly expressed in ovarian tumor tissue
CC  (and some also in testis and breast cancer tissue). The products of the
CC  invention can be used for gene therapy. (A) are used (i) for recombinant
CC  expression of polypeptides (B) and (ii) to isolate complete genes. (B)
CC  are used (i) to identify agents suitable for treatment of ovarian cancer;
CC  (ii) directly for treating this form of cancer (including expression from
CC  gene therapy vectors) and (iii) for generation of specific antibodies.
CC  (A) are identified by assembling ESTs (expressed sequence tags) from a
CC  particular tissue type before comparison of expression patterns. This
CC  allows a significantly longer fragment of the gene to be revealed, so
CC  should reduce the number of failures associated with the fact that ESTs
CC  from different libraries may represent different parts of the same
CC  unknown gene, distorting the estimated frequency of occurrence in a
CC  particular tissue. AAZ77450-277572 represent the human ovarian tumor cDNA
CC  library derived EST fragments described in the method of the invention
CC  and encode the protein fragments represented in AAY76505-Y76638
XX
SQ  Sequence 599 BP; 142 A; 153 C; 154 G; 150 T; 0 U; 0 Other;
XX

Alignment Scores:
Pred. No.:             2.11e-58        Length:        599
Score:                 762.00         Matches:       132
Percent Similarity:    100.00%        Conservative:  0
Best Local Similarity: 100.00%        Mismatches:    0
Query Match:           100.00%        Indels:        0
DB:                                   Gaps:          0

US-10-613-105-2 (1-132) x AAZ77546 (1-599)

Qy    1 MetLysSerSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro  20
Db   14 ATGAAGTCCAGCGGCCCTTCCCCTTCCTGCGTGCTGGCTCTGGGAACTCTGGCACCT  73
Qy   21 TrpAlaValGluGlyLysSerGlyLysSerPheLysAlaGlyLysValCysProLysSer  40
Db   74 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCGTCCTCCTAAGAAATCT 133
Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
Db  134 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 193
Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
Db  194 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 253
Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyLysCysLeuMetLeuAsn 100
Db  254 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 313
Qy  101 ProProAsnPheCysGluMetAspGlyLysCysLysArgAspLeuLysCysCysMetGly 120
Db  314 CCCCCAAATTTCTGTGAGATGGATGGCAAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 373
```

```
Qy    21  TrpAlaValGluGlySerGlyLysSerGlyLysPheLysAlaGlyValCysProProLysLysSer  40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116  TGGGCTGTGGAAGGCTCTGGAAGTCCTTCAAAGCTCGAGTCTGTCCTCCTAAGAAATCT        175

Qy    41  AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerGlnSerAspTrpGlnCysProGlyLys  60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   176  GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTGACTGGCAGTGTCCAGGGAAG        235

Qy    61  LysArgCysCysProAspThrCysGlyIleLeuCysGlyLeuAspProValAspThrProAsn  80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   236  AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCTGGTGTCTGTTGACACCCCAAAC        295

Qy    81  ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn  100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   296  CCAACAGGAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTGATGCTTAAC        355

Qy   101  ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysAspCysMetGly  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   356  CCCCCAATTTCTGTGATGATGGCCAGTGCAAGAAGCGTGACTTGAAGTGTGACTGCATGGGC        415

Qy   121  MetCysGlyLysSerCysValSerProValLysAla  132
          ||||||||||||||||||||||||||||||||||||
Db   416  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT        451
```

RESULT 27
ADL62375
ID   ADL62375 standard; DNA; 1084 BP.
XX
AC   ADL62375;
XX
DT   20-MAY-2004 (first entry)
XX
DE   Human ovarian cancer DNA marker #20587.
XX
KW   Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX
OS   Homo sapiens.
XX
PN   WO200170979-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US009126.
XX
PR   21-MAR-2000; 2000US-0191031P.
PR   25-MAY-2000; 2000US-0207124P.
PR   15-JUN-2000; 2000US-0211940P.
PR   07-JUL-2000; 2000US-0216820P.
PR   25-JUL-2000; 2000US-0220661P.
PR   21-DEC-2000; 2000US-0257672P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Lee J, Lillie J;
XX
DR   WPI; 2001-611502/70.
XX
PT   Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
PT   cancer cells as compared to their normal non-cancerous ovarian cells are
PT   used to characterize stage, grade, histological type of ovarian cancer.
XX
PS   Disclosure; SEQ ID NO 20587; 106pp; English.
XX
CC   The invention relates to nucleic acid markers which are overexpressed in
CC   ovarian cancer cells as compared to their expression in normal (i.e. non-
CC   cancerous) ovarian cells. The invention also relates to polypeptides
CC   encoded by the markers, antibodies that selectively bind to the
CC   polypeptides, a method of inhibiting ovarian cancer in a patient at risk
CC   of developing ovarian cancer involving inhibiting expression of a gene
CC   corresponding to a marker of the invention and a method of treating a
CC   patient afflicted with ovarian cancer comprising providing to cells of
CC   the patient an antisense oligonucleotide complementary to a marker of the

```
Qy   121  MetCysGlyLysSerCysValSerProValLysAla  132
          ||||||||||||||||||||||||||||||||||||
Db   374  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT        409
```

RESULT 26
AAH57451
ID   AAH57451 standard; cDNA; 636 BP.
XX
AC   AAH57451;
XX
DT   10-SEP-2001 (first entry)
XX
DE   Human lung cell specific cDNA sequence SEQ ID NO:291.
XX
KW   Human; tissue specific; diagnosis; brain; heart; skeletal muscle; lung;
KW   liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW   metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW   neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX
OS   Homo sapiens.
XX
PN   WO200132927-A2.
XX
PD   10-MAY-2001.
XX
PF   02-NOV-2000; 2000WO-US030396.
XX
PR   04-NOV-1999; 99US-0163508P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Sornasse T, Seilhamer JJ, Watson GA;
XX
DR   WPI; 2001-291057/30.
XX
PT   New cell and tissue specific polynucleotides useful for diagnosis,
PT   prognosis or monitoring of treatments for disorders where the gene is
PT   associated with a cancer, immunopathology or neuropathology.
XX
PS   Claim 1; Page 217; 327pp; English.
XX
CC   AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC   sequences (II). (I) can have cytostatic, immunomodulatory and
CC   neuroprotective activities, and can be used in gene therapy. (I) and
CC   proteins (II) encoded by then are used in high throughput screening
CC   assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC   mimetics, peptides, proteins, agonists, antagonists, antibodies or their
CC   fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical
CC   agents. Expression of (I) in a sample indicates the differentiation of
CC   embryonic stem cells into a tissue selected from brain, heart, kidney,
CC   liver, lung, skeletal muscle or pancreatic tissues. (I) and (II) are used
CC   to produce an expression profile that defines a metabolic or
CC   developmental process, treatment, condition, disease or disorder. The
CC   gene profile can be used for diagnosis, prognosis or monitoring of
CC   treatments and for investigating a predisposition to a disorder where the
CC   gene is associated with a cancer, immunopathology or neuropathology
XX
SQ   Sequence 636 BP; 160 A; 159 C; 163 G; 154 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   2.26e-58   Length: 636
Score: 762.00   Matches: 132
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 4   Gaps: 0

US-10-613-105-2 (1-132) x AAH57451 (1-636)

Qy   1 MetLysSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   56 ATGAAGTCCAGGGAGCCCTCTTCCCCTTCCTCCTGGTGCTGTTGCCCTGGAACTCTGGCACCT 115
```

```
CC   invention. The markers are useful for assessing if a patient is afflicted
CC   with ovarian cancer, which involves comparing the level of expression of
CC   a marker in a patient sample and a normal level of expression of the
CC   marker in a control non-ovarian cancer sample. A difference between the
CC   expression levels indicates ovarian cancer. The level of expression of a
CC   marker corresponds to a secreted protein or to a transcribed
CC   polynucleotide or its portion. The level of expression of the marker is
CC   assessed by detecting the presence in the sample, a protein or protein
CC   fragment corresponding to the marker. The presence of protein or protein
CC   fragment is detected using an antibody that specifically binds with the
CC   protein or protein fragment. Alternatively, the level of expression of
CC   the marker is assessed by detecting the presence of a transcribed
CC   polynucleotide which anneals with the marker or anneals with a portion of
CC   the polynucleotide comprising the marker, under stringent conditions. The
CC   marker is also used for monitoring the progression of ovarian cancer in a
CC   patient which involves detecting expression of the marker in a patient
CC   sample at a first point in time, repeating the method at a subsequent
CC   time and comparing the level of expression. The method is carried out
CC   using an ovarian tissue sample. A composition comprising a marker,
CC   polypeptide or antibody of the invention is used to treat ovarian cancer.
CC   This sequence represents a human ovarian cancer DNA marker of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1084 BP; 267 A; 264 C; 329 G; 220 T; 0 U; 4 Other;
```

```
Alignment Scores:
Pred. No.:                          Length:        1084
Score:                4.2e-58       Matches:       132
Percent Similarity:   762.00        Conservative:  0
Best Local Similarity: 100.00%      Mismatches:    0
Query Match:          100.00%       Indels:        0
DB:                                 Gaps:          5
```

```
US-10-613-105-2 (1-132) x ADL62375 (1-1084)

Qy    1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro     20
Db   69 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTGGCCCTGGGAACTCTGGCACCT    128

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysSerAlaAlaGlyTyrValCysProProLysLysSer  40
Db  129 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGTGGCAGCTGGTTACGTGTGTCCTCCTAAGAAATCT  188

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys     60
Db  189 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG    248

Qy   61 LysArgCysCysProAspThrCysGlyIleLeuLysCysGlyLeuAspProValAspThrProAsn  80
Db  249 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC    308

Qy   81 ProThrArgArgLysProGlyLysProValThrTyrTyrGlnCysLeuMetLeuAsn       100
Db  309 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC    368

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly  120
Db  369 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGGTGACTTGAAGTGTTGCATGGGC    428

Qy  121 MetCysGlyLysLysSerCysValSerProValLysSerAla                      132
Db  429 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT                            464
```

```
RESULT 28
ADL57146
ID   ADL57146 standard; DNA; 594 BP.
XX
AC   ADL57146;
XX
DT   03-JUN-2004  (first entry)
XX
```

```
DE   Human NOV8d gene SEQ ID NO:91.
XX
KW   ds; gene; human; antidiabetic; anorectic; cardiant; hypotensive;
KW   antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
KW   protozoacide; nootropic; neuroprotective; antiparkinsonian;
KW   anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
KW   dermatological; antiasthmatic; antilipaemic; gene therapy;
KW   fibroblast growth factor receptor 4; FGFR4;
KW   complement factor I precursor; matrix metalloproteinase-15 precursor;
KW   MDC3; T-lymphocyte surface antigen Ly-9 precursor;
KW   fibroblast growth factor-21; FGF-21;
KW   alpha-2 macroglobulin-like polypeptide variant;
KW   antileukoproteinase 1 precursor; LiV-1; nuclear hormone receptor NOR-1;
KW   transmembrane protein-like; beta-neoendorphin-dynorphin precursor.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   CDS            19..417
FT                  /*tag=  a
XX
PN   WO2004022723-A2.
XX
PD   18-MAR-2004.
XX
PF   09-SEP-2003; 2003WO-US028141.
XX
PR   09-SEP-2002; 2002US-0409145P.
PR   10-SEP-2002; 2002US-0409544P.
PR   12-SEP-2002; 2002US-0410320P.
PR   16-SEP-2002; 2002US-0411060P.
PR   23-SEP-2002; 2002US-0412766P.
PR   23-SEP-2002; 2002US-0412825P.
PR   24-SEP-2002; 2002US-0412767P.
PR   25-SEP-2002; 2002US-0413342P.
PR   30-SEP-2002; 2002US-0414832P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Zhong M, Guo X, Anderson DW, Ort T, Padigaru M, Rieger DK;
XX
DR   WPI; 2004-315567/29.
DR   P-PSDB; ADL57147.
XX
PT   New isolated NOVX polypeptides and polynucleotides, useful for
PT   preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT   osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT   asthma, or infections.
XX
PS   Claim 17; SEQ ID NO 91; 214pp; English.
XX
CC   The invention relates to a novel isolated polypeptide (NOVX) comprising a
CC   mature form of any of the 37 amino acid sequences fully defined in the
CC   specification. A polypeptide of the invention has antidiabetic,
CC   anorectic, cardiant, hypotensive, antiarteriosclerotic, anorectic,
CC   virucide, antibacterial, fungicide, protozoacide, nootropic,
CC   neuroprotective, antiparkinsonian, anticonvulsant, osteopathic,
CC   antiarthritic, antiinflammatory, dermatological, antiasthmatic, and
CC   antilipaemic activity. A polynucleotide of the invention may have a use
CC   in gene therapy. The polypeptides, nucleic acid molecules and antibodies
CC   are useful in the manufacture of a medicament for treating a syndrome
CC   associated with a human disease, preferably a NOVX-associated disorder.
CC   The nucleic acid molecules, polypeptides and antibodies are useful for
CC   treating, preventing or diagnosing diseases such as metabolic disorders,
CC   diabetes, obesity, infectious diseases (viral, bacterial, fungal,
CC   helminthic, and protozoal) anorexia, cancer, cardiovascular diseases
CC   (hypertension, atherosclerosis), neurodegenerative disorders, Alzheimer's
CC   disease, Parkinson's disease, epilepsy, immune disorders
CC   (osteoarthritis), haematopoietic disorders, inflammatory skin disorders,
CC   asthma, and various dyslipidaemias. The nucleic acids and polypeptides
CC   may also be used as targets for the identification of small molecules
CC   that modulate or inhibit e.g. neurogenesis, cell differentiation, cell
CC   proliferation, haematopoiesis, wound healing and angiogenesis, in gene
```

```
CC      therapy, in generation of antibodies that bind immunospecifically to NOVX
CC      substances for use in therapeutic or diagnostic methods. The nucleic
CC      acids are further used as hybridisation probes, in chromosome mapping,
CC      tissue typing, preventive medicine, and pharmacogenomics. The NOVX
CC      polypeptides of the invention show homology to certain known human
CC      proteins: NOV1a-1t show homology to fibroblast growth factor receptor 4
CC      (FGFR4); NOV2a shows homology to complement factor I precursor; NOV3a
CC      shows homology to matrix metalloproteinase-15 precursor; NOV4a shows
CC      homology to MDC3; NOV5a-5c show homology to T-lymphocyte surface antigen
CC      Ly-9 precursor; NOV6a-6m show homology to fibroblast growth factor-21
CC      (FGF-21); NOV7a-7c show homology to alpha-2 macroglobulin-like
CC      polypeptide variant; NOV8a-8g show homology to antileukoproteinase 1
CC      precursor; NOV9a-9i show homology to LIV-1 protein; NOV10a shows homology
CC      to nuclear hormone receptor NOR-1; NOV11a-11j show homology to
CC      transmembrane protein-like; NOV12a-12c show homology to beta-neoendorphin
CC      -dynorphin precursor. The present sequence encodes a NOVX polypeptide of
CC      the invention.
XX
SQ      Sequence 594 BP; 132 A; 157 C; 155 G; 150 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                                Length:          594
Score:               7.07e-58             Matches:         131
Percent Similarity:  99.24%               Conservative:      0
Best Local Similarity: 99.24%             Mismatches:        1
Query Match:         99.21%               Indels:            0
DB:                  12                   Gaps:              0

US-10-613-105-2 (1-132) x ADL57146 (1-594)

Qy      1  MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     19  ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTTCTCGCTCTGGGAACCCTGGCCACCT 78

Qy     21  TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     79  TGGGCTGTGGAAGGCTCGGAAAGTCCTTCAAAGTGGAGTCTGTCCTCCTAAGAAATCT 138

Qy     41  AlaGlnCysLeuArgTyrLysLysProGluCysProGlnSerAspTrpGlnCysProGlyLys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    139  GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy     61  LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    199  AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTAGACCCTGTGGACACCCCAAAC 258

Qy     81  ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    259  CCAACAAGGACGAAGCCTGGAAGTGCCCAGTGACTTATGGCCAATGTTCCATGCTTAAC 318

Qy    101  ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    319  CCCCCAAATTTCTGTGAAATGGATGGCCAAGGTGACTTGAAGTGTTGCATGGGC 378

Qy    121  MetCysGlyLysSerCysValSerProValLysAla 132
           |||||||||||||||||||||||||||||||||
Db    379  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414

RESULT 29
ADL57144
ID      ADL57144 standard; DNA; 594 BP.
XX
AC      ADL57144;
XX
DT      03-JUN-2004  (first entry)
XX
DE      Human NOV8c gene SEQ ID NO:89.
XX
KW      ds; gene; human; antidiabetic; anorectic; cardiant; hypotensive;
KW      antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
KW      protozoacide; nootropic; neuroprotective; antiparkinsonian;
KW      anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
KW      dermatological; antiasthmatic; antilipaemic; gene therapy;
```

```
KW      fibroblast growth factor receptor 4; FGFR4;
KW      complement factor I precursor; matrix metalloproteinase-15 precursor;
KW      MDC3; T-lymphocyte surface antigen Ly-9 precursor;
KW      fibroblast growth factor-21; FGF-21;
KW      alpha-2 macroglobulin-like polypeptide variant;
KW      antileukoproteinase 1 precursor; LIV-1; nuclear hormone receptor NOR-1;
KW      transmembrane protein-like; beta-neoendorphin-dynorphin precursor.
XX
OS      Homo sapiens.
XX
FH      Key             Location/Qualifiers
FT      CDS             19..417
FT                      /*tag= a
XX
PN      WO200402273-A2.
XX
PD      18-MAR-2004.
XX
PF      09-SEP-2003; 2003WO-US028141.
XX
PR      09-SEP-2002; 2002US-0409145P.
PR      10-SEP-2002; 2002US-0409544P.
PR      12-SEP-2002; 2002US-0410320P.
PR      16-SEP-2002; 2002US-0411060P.
PR      23-SEP-2002; 2002US-0412766P.
PR      23-SEP-2002; 2002US-0412825P.
PR      24-SEP-2002; 2002US-0412767P.
PR      25-SEP-2002; 2002US-0413342P.
PR      30-SEP-2002; 2002US-0414832P.
XX
PA      (CURA-) CURAGEN CORP.
XX
PI      Zhong M, Guo X, Anderson DW, Ort T, Padigaru M, Rieger DK;
XX
DR      WPI; 2004-315567/29.
DR      P-PSDB; ADL57145.
XX
PT      New isolated NOVX polypeptides and polynucleotides, useful for
PT      preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT      osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT      asthma, or infections.
XX
PS      Claim 17; SEQ ID NO 89; 214pp; English.
XX
CC      The invention relates to a novel isolated polypeptide (NOVX) comprising a
CC      mature form of any of the 37 amino acid sequences fully defined in the
CC      specification. A polypeptide of the invention has antidiabetic,
CC      anorectic, cardiant, hypotensive, antiarteriosclerotic, anorectic,
CC      virucide, antibacterial, fungicide, protozoacide, nootropic,
CC      neuroprotective, antiparkinsonian, anticonvulsant, osteopathic,
CC      antiarthritic, antiinflammatory, dermatological, antiasthmatic, and
CC      antilipaemic activity. A polynucleotide of the invention may have a use
CC      in gene therapy. The polypeptides, nucleic acid molecules and antibodies
CC      are useful in the manufacture of a medicament for treating a syndrome
CC      associated with a human disease, preferably a NOVX-associated disorder.
CC      The nucleic acid molecules, polypeptides and antibodies are useful for
CC      treating, preventing or diagnosing diseases such as metabolic disorders,
CC      diabetes, obesity, infectious diseases (viral, bacterial, fungal,
CC      helminthic, and protozoal), anorexia, cancer, cardiovascular diseases
CC      (hypertension, atherosclerosis), neurodegenerative disorders, Alzheimer's
CC      disease, Parkinson's disease, epilepsy, immune disorders
CC      (osteoarthritie), haematopoietic disorders, inflammatory skin disorders,
CC      asthma, and various dyslipidaemias. The nucleic acids and polypeptides
CC      may also be used as targets for the identification of small molecules
CC      that modulate or inhibit e.g. neurogenesis, cell differentiation, cell
CC      proliferation, haematopoiesis, wound healing and angiogenesis, in gene
CC      therapy, in generation of antibodies that bind immunospecifically to NOVX
CC      substances for use in therapeutic or diagnostic methods. The nucleic
CC      acids are further used as hybridisation probes, in chromosome mapping,
CC      tissue typing, preventive medicine, and pharmacogenomics. The NOVX
CC      polypeptides of the invention show homology to certain known human
CC      proteins: NOV1a-1t show homology to fibroblast growth factor receptor 4
CC      (FGFR4); NOV2a shows homology to complement factor I precursor; NOV3a
```

```
CC    shows homology to matrix metalloproteinase-15 precursor; NOV4a shows
CC    homology to MDC3; NOV5a-5c show homology to T-lymphocyte surface antigen
CC    Ly-9 precursor; NOV6a-6m show homology to fibroblast growth factor-21
CC    (FGF-21); NOV7a-7c show homology to alpha-2 macroglobulin-like
CC    polypeptide variant; NOV8a-8g show homology to antileukoproteinase 1
CC    precursor; NOV9a-9i show homology to LIV-1 protein; NOV10a shows homology
CC    to nuclear hormone receptor NOR-1; NOV11a-11j show homology to
CC    transmembrane protein-like; NOV12a-12c show homology to beta-neoendorphin
CC    -dynorphin precursor. The present sequence encodes a NOVX polypeptide of
CC    the invention.
XX
SQ    Sequence 594 BP; 132 A; 157 C; 155 G; 150 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                          Length:         594
Score:               8.67e-58       Matches:        131
Percent Similarity:  755.00         Conservative:     0
Best Local Similarity: 99.24%       Mismatches:       0
Query Match:         99.08%         Indels:           0
DB:                  12             Gaps:             0

US-10-613-105-2 (1-132) x ADL57144 (1-594)

Qy      1 MetLysSerSerGlyLeuPheProPheLeuValLeuAlaLeuAlaGlyThrLeuAlaPro 20
Db     19 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTCGCCCTGGCAACTCTGGCACCT 78

Qy     21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyLysValCysProProLysLysSer 40
Db     79 TGGGCTGTGGAGGGCCTCGGAAAGTCCTTCAAAGCTGGAGTGCCTCCTCCTAAGAAATCT 138

Qy     41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnSerProGlyLys 60
Db    139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy     61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
Db    199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy     81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
Db    259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy    101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
Db    319 CCCCCAAATTTCTGTGAAATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy    121 MetCysGlyLysSerCysValSerProValLysAla 132
Db    379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414

RESULT 30
ADL57140
ID    ADL57140 standard; DNA; 594 BP.
XX
AC    ADL57140;
XX
DT    03-JUN-2004 (first entry)
XX
DE    Human NOV8a gene SEQ ID NO:85.
XX
KW    ds; gene; human; antidiabetic; anorectic; cardiant; hypotensive;
KW    antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
KW    protozoacide; nootropic; neuroprotective; antiparkinsonian;
KW    anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
KW    dermatological; antiasthmatic; antilipaemic; gene therapy;
KW    fibroblast growth factor receptor 4; FGFR4;
KW    complement factor I precursor; matrix metalloproteinase-15 precursor;
KW    MDC3; T-lymphocyte surface antigen Ly-9 precursor;
KW    fibroblast growth factor-21; FGF-21;
KW    alpha-2 macroglobulin-like polypeptide variant;
KW    antileukoproteinase 1 precursor; LIV-1; nuclear hormone receptor NOR-1;
KW    transmembrane protein-like; beta-neoendorphin-dynorphin precursor.
```

```
XX
OS    Homo sapiens.
XX
FH    Key                  Location/Qualifiers
FT    CDS                  1..417
FT                         /*tag= a
XX
PN    WO2004022723-A2.
XX
PD    18-MAR-2004.
XX
PF    09-SEP-2003; 2003WO-US028141.
XX
PR    09-SEP-2002; 2002US-0409145P.
PR    10-SEP-2002; 2002US-0409544P.
PR    12-SEP-2002; 2002US-0410320P.
PR    16-SEP-2002; 2002US-0411060P.
PR    23-SEP-2002; 2002US-0412766P.
PR    23-SEP-2002; 2002US-0412825P.
PR    24-SEP-2002; 2002US-0412767P.
PR    25-SEP-2002; 2002US-0413342P.
PR    30-SEP-2002; 2002US-0414832P.
XX
PA    (CURA-) CURAGEN CORP.
XX
PI    Zhong M,  Guo X,  Anderson DW,  Ort T,  Padigaru M,  Rieger DK;
XX
DR    WPI; 2004-315567/29.
DR    P-PSDB; ADL57141.
XX
PT    New isolated NOVX polypeptides and polynucleotides, useful for
PT    preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT    osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT    asthma, or infections.
XX
PS    Claim 17; SEQ ID NO 85; 214pp; English.
XX
CC    The invention relates to a novel isolated polypeptide (NOVX) comprising a
CC    mature form of any of the 37 amino acid sequences fully defined in the
CC    specification. A polypeptide of the invention has antidiabetic,
CC    anorectic, cardiant, hypotensive, antiarteriosclerotic, anorectic,
CC    virucide, antibacterial, fungicide, protozoacide, nootropic,
CC    neuroprotective, antiparkinsonian, anticonvulsant, osteopathic,
CC    antiarthritic, antiinflammatory, dermatological, antiasthmatic, and
CC    antilipaemic activity. A polynucleotide of the invention may have a use
CC    in gene therapy. The polypeptides, nucleic acid molecules and antibodies
CC    are useful in the manufacture of a medicament for treating a syndrome
CC    associated with a human disease, preferably a NOVX-associated disorder.
CC    The nucleic acid molecules, polypeptides and antibodies are useful for
CC    treating, preventing or diagnosing diseases such as metabolic disorders,
CC    diabetes, obesity, infectious diseases (viral, bacterial, fungal,
CC    helminthic, and protozoal), anorexia, cancer, cardiovascular diseases
CC    (hypertension, atherosclerosis), neurodegenerative disorders, Alzheimer's
CC    disease, Parkinson's disease, epilepsy, immune disorders
CC    (osteoarthritis), haematopoietic disorders, inflammatory skin disorders,
CC    asthma, and various dyslipidaemias. The nucleic acids and polypeptides
CC    may also be used as targets for the identification of small molecules
CC    that modulate or inhibit e.g. neurogenesis, cell differentiation, cell
CC    proliferation, haematopoiesis, wound healing and angiogenesis, in gene
CC    therapy, in generation of antibodies that bind immunospecifically to NOVX
CC    substances for use in therapeutic or diagnostic methods. The nucleic
CC    acids are further used as hybridisation probes, in chromosome mapping,
CC    tissue typing, preventive medicine, and pharmacogenomics. The NOVX
CC    polypeptides of the invention show homology to certain known human
CC    proteins: NOV1a-1t show homology to fibroblast growth factor receptor 4
CC    (FGFR4); NOV2a shows homology to complement factor I precursor; NOV3a
CC    shows homology to matrix metalloproteinase-15 precursor; NOV4a shows
CC    homology to MDC3; NOV5a-5c show homology to T-lymphocyte surface antigen
CC    Ly-9 precursor; NOV6a-6m show homology to fibroblast growth factor-21
CC    (FGF-21); NOV7a-7c show homology to alpha-2 macroglobulin-like
CC    polypeptide variant; NOV8a-8g show homology to antileukoproteinase 1
CC    precursor; NOV9a-9i show homology to LIV-1 protein; NOV10a shows homology
CC    to nuclear hormone receptor NOR-1; NOV11a-11j show homology to
```

Wed Oct 27 14:05:49 2004   us-10-613-105-2.rng   Page 24

CC   transmembrane protein-like; NOV12a-12c show homology to beta-neoendorphin
CC   -dynorphin precursor. The present sequence encodes a NOVX polypeptide of
CC   the invention.
XX
SQ   Sequence 594 BP; 132 A; 157 C; 155 G; 150 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:		Length:	594
Score:	1.06e-57	Matches:	131
Percent Similarity:	99.24%	Conservative:	0
Best Local Similarity:	99.24%	Mismatches:	1
Query Match:	98.95%	Indels:	0
DB:		Gaps:	0

12

US-10-613-105-2 (1-132) x ADL57140 (1-594)

```
Qy    1 MetLysSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro      20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 ATGAAGTCCAGCGGCCTCTCCCCCTTCCTGGTGCTGCTGGCCCTGGGAACTCTGGCACCT   78

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysSer      40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT  138

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG  198

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn    80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC  258

Qy   81 ProThrArgArgLysProGlyLysProValThrTyrGlyGlnCysLeuMetLeuAsn     100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259 CCAACAAGAAGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC  318

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC  378

Qy  121 MetCysGlyLysLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||||||||||||||
Db  379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414
```

RESULT 31
ADL57142

ID   ADL57142 standard; DNA; 594 BP.
XX
AC   ADL57142;
XX
DT   03-JUN-2004 (first entry)
XX
DE   Human NOV8b gene SEQ ID NO:87.
XX
KW   ds; gene; human; antidiabetic; anorexic; cardiant; hypotensive;
KW   antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
KW   protozoacide; nootropic; neuroprotective; antiparkinsonian;
KW   anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
KW   dermatological; antiasthmatic; antilipaemic; gene therapy;
KW   fibroblast growth factor receptor 4; FGFR4;
KW   complement factor I precursor; matrix metalloproteinase-15 precursor;
KW   MDC3; T-lymphocyte surface antigen Ly-9 precursor;
KW   fibroblast growth factor-21; FGF-21;
KW   alpha-2 macroglobulin-like polypeptide variant;
KW   antileukoproteinase 1 precursor; LIV-1; nuclear hormone receptor NOR-1;
KW   transmembrane protein-like; beta-neoendorphin-dynorphin precursor.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   19..417
FT   /*tag= a
XX

PN   WO200402723-A2.
XX
PD   18-MAR-2004.
XX
PF   09-SEP-2003; 2003WO-US028141.
XX
PR   09-SEP-2002; 2002US-0409145P.
PR   10-SEP-2002; 2002US-0409544P.
PR   12-SEP-2002; 2002US-0410320P.
PR   16-SEP-2002; 2002US-0411060P.
PR   23-SEP-2002; 2002US-0412766P.
PR   23-SEP-2002; 2002US-0412825P.
PR   24-SEP-2002; 2002US-0412767P.
PR   25-SEP-2002; 2002US-0413342P.
PR   30-SEP-2002; 2002US-0414832P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Zhong M, Guo X, Anderson DW, Ort T, Padigaru M, Rieger DK;
XX
DR   WPI; 2004-315567/29.
DR   P-PSDB; ADL57143.
XX
PT   New isolated NOVX polypeptides and polynucleotides, useful for
PT   preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT   osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT   asthma, or infections.
XX
PS   Claim 17; SEQ ID NO 87; 214pp; English.
XX
CC   The invention relates to a novel isolated polypeptide (NOVX) comprising a
CC   mature form of any of the 37 amino acid sequences fully defined in the
CC   specification. A polypeptide of the invention has antidiabetic,
CC   anorectic, cardiant, hypotensive, antiarteriosclerotic, anorectic,
CC   virucide, antibacterial, fungicide, protozoacide, nootropic,
CC   neuroprotective, antiparkinsonian, anticonvulsant, osteopathic,
CC   antiarthritic, antiinflammatory, dermatological, antiasthmatic, and
CC   antilipaemic activity. A polynucleotide of the invention may have a use
CC   in gene therapy. The polypeptides, nucleic acid molecules and antibodies
CC   are useful in the manufacture of a medicament for treating a syndrome
CC   associated with a human disease, preferably a NOVX-associated disorder.
CC   The nucleic acid molecules, polypeptides and antibodies are useful for
CC   treating, preventing or diagnosing diseases such as metabolic disorders,
CC   diabetes, obesity, infectious diseases (viral, bacterial, fungal,
CC   helminthic, and protozoal), anorexia, cancer, cardiovascular diseases
CC   (hypertension, atherosclerosis), neurodegenerative disorders, Alzheimer's
CC   disease, Parkinson's disease, epilepsy, immune disorders
CC   (osteoarthritis), haematopoietic disorders, inflammatory skin disorders,
CC   asthma, and various dyslipidaemias. The nucleic acids and polypeptides
CC   may also be used as targets for the identification of small molecules
CC   that modulate or inhibit e.g. neurogenesis, cell differentiation, cell
CC   proliferation, haematopoiesis, wound healing and angiogenesis, in gene
CC   therapy, in generation of antibodies that bind immunospecifically to NOVX
CC   substances for use in therapeutic or diagnostic methods. The nucleic
CC   acids are further used as hybridisation probes, in chromosome mapping,
CC   tissue typing, preventive medicine, and pharmacogenomics. The NOVX
CC   polypeptides of the invention show homology to certain known human
CC   proteins: NOV1a-1t show homology to fibroblast growth factor receptor 4
CC   (FGFR4); NOV2a shows homology to complement factor I precursor; NOV3a
CC   shows homology to matrix metalloproteinase-15 precursor; NOV4a shows
CC   homology to MDC3; NOV5a-5c show homology to T-lymphocyte surface antigen
CC   Ly-9 precursor; NOV6a-6m show homology to fibroblast growth factor-21
CC   (FGF-21); NOV7a-7c show homology to alpha-2 macroglobulin-like
CC   polypeptide variant; NOV8a-8g show homology to antileukoproteinase 1
CC   precursor; NOV9a-9i show homology to LIV-1 protein; NOV10a shows homology
CC   to nuclear hormone receptor NOR-1; NOV11a-11j show homology to
CC   transmembrane protein-like; NOV12a-12c show homology to beta-neoendorphin
CC   -dynorphin precursor. The present sequence encodes a NOVX polypeptide of
CC   the invention.
XX
SQ   Sequence 594 BP; 132 A; 157 C; 155 G; 150 T; 0 U; 0 Other;

Alignment Scores:

```
Pred. No.:              2.4e-57      Length:          594
Score:                  750.00       Matches:         131
Percent Similarity:     99.24%       Conservative:    0
Best Local Similarity:  99.24%       Mismatches:      1
Query Match:            98.43%       Indels:          0
DB:                                  Gaps:            0
```

US-10-613-105-2 (1-132) x ADL57142 (1-594)

```
Qy      1 MetLysSerSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     19 ATGAAGTCCAGGGGCCTCTTCCCCTTCCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCT 78

Qy     21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 138

Qy     41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
            ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db    139 GCCCAGTGCCTTAGATACAAGAAACCTGAGCGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy     61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy     81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyTyrGlyLeuMetLeuAsn 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCTATGGTTTGATGCTTAAC 318

Qy    101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy    121 MetCysGlyLysSerCysValSerProValLysAla 132
            ||||||||||||||||||||||||||||||
Db    379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414
```

RESULT 32
ADL57148

```
ID   ADL57148 standard; DNA; 594 BP.
XX
AC   ADL57148;
XX
DT   03-JUN-2004  (first entry)
XX
DE   Human NOV8e gene SEQ ID NO:93.
XX
KW   ds; gene; human; antidiabetic; anorectic; cardiant; hypotensive;
KW   antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
KW   procozoacide; nootropic; neuroprotective; antiparkinsonian;
KW   anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
KW   dermatological; antiasthmatic; antilipaemic; gene therapy;
KW   fibroblast growth factor receptor 4; FGFR4;
KW   complement factor I precursor; matrix metalloproteinase-15 precursor;
KW   MDC3; T-lymphocyte surface antigen Ly-9 precursor;
KW   fibroblast growth factor-21; FGF-21;
KW   alpha-2 macroglobulin-like polypeptide variant;
KW   antileukoproteinase 1 precursor; LIV-1; nuclear hormone receptor NOR-1;
KW   transmembrane protein-like; beta-neoendorphin-dynorphin precursor.
XX
OS   Homo sapiens.
XX
FH                    Location/Qualifiers
FT   CDS              19..417
FT                    /*tag= a
XX
PN   WO2004022723-A2.
XX
PD   18-MAR-2004.
XX
PF   09-SEP-2003; 2003WO-US028141.
XX
PR   09-SEP-2002; 2002US-0409145P.
```

```
PR   10-SEP-2002;  2002US-0409544P.
PR   12-SEP-2002;  2002US-0410320P.
PR   16-SEP-2002;  2002US-0411060P.
PR   23-SEP-2002;  2002US-0412766P.
PR   23-SEP-2002;  2002US-0412825P.
PR   24-SEP-2002;  2002US-0412767P.
PR   25-SEP-2002;  2002US-0413342P.
PR   30-SEP-2002;  2002US-0414832P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Zhong M, Guo X, Anderson DW, Ort T, Padigaru M, Rieger DK;
XX
DR   WPI; 2004-315567/29.
DR   P-PSDB; ADL57149.
XX
PT   New isolated NOVX polypeptides and polynucleotides, useful for
PT   preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT   osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT   asthma, or infections.
XX
PS   Claim 17; SEQ ID NO 93; 214pp; English.
XX
CC   The invention relates to a novel isolated polypeptide (NOVX) comprising a
CC   mature form of any of the 37 amino acid sequences fully defined in the
CC   specification. A polypeptide of the invention has antidiabetic,
CC   anorectic, cardiant, hypotensive, antiarteriosclerotic, anorectic,
CC   virucide, antibacterial, fungicide, procozoacide, nootropic,
CC   neuroprotective, antiparkinsonian, anticonvulsant, osteopathic,
CC   antiarthritic, antiinflammatory, dermatological, antiasthmatic, and
CC   antilipaemic activity. A polynucleotide of the invention may have a use
CC   in gene therapy. The polypeptides, nucleic acid molecules and antibodies
CC   are useful in the manufacture of a medicament for treating a syndrome
CC   associated with a human disease, preferably a NOVX-associated disorder.
CC   The nucleic acid molecules, polypeptides and antibodies are useful for
CC   treating, preventing or diagnosing diseases such as metabolic disorders,
CC   diabetes, obesity, infectious diseases (viral, bacterial, fungal,
CC   helminthic, and protozoal), anorexia, cancer, cardiovascular diseases
CC   (hypertension, atherosclerosis), neurodegenerative disorders, Alzheimer's
CC   disease, Parkinson's disease, epilepsy, immune disorders
CC   (osteoarthritis), haematopoietic disorders, inflammatory skin disorders,
CC   asthma, and various dyslipidaemias. The nucleic acids and polypeptides
CC   may also be used as targets for the identification of small molecules
CC   that modulate or inhibit e.g. neurogenesis, cell differentiation, cell
CC   proliferation, haematopoiesis, wound healing and angiogenesis, in gene
CC   therapy, in generation of antibodies that bind immunospecifically to NOVX
CC   substances for use in therapeutic or diagnostic methods. The nucleic
CC   acids are further used as hybridisation probes, in chromosome mapping,
CC   tissue typing, preventive medicine, and pharmacogenomics. The NOVX
CC   polypeptides of the invention show homology to certain known human
CC   proteins: NOV1a-1t show homology to fibroblast growth factor receptor 4
CC   (FGFR4); NOV2a shows homology to complement factor I precursor; NOV3a
CC   shows homology to matrix metalloproteinase-15 precursor; NOV4a shows
CC   homology to MDC3; NOV5a-5c show homology to T-lymphocyte surface antigen
CC   Ly-9 precursor; NOV6a-6m show homology to fibroblast growth factor-21
CC   (FGF-21); NOV7a-7c show homology to alpha-2 macroglobulin-like
CC   polypeptide variant; NOV8a-8g show homology to antileukoproteinase 1
CC   precursor; NOV9a-9i show homology to LIV-1 protein; NOV10a shows homology
CC   to nuclear hormone receptor NOR-1; NOV11a-11j show homology to
CC   transmembrane protein-like; NOV12a-12c show homology to beta-neoendorphin
CC   -dynorphin precursor. The present sequence encodes a NOVX polypeptide of
CC   the invention.
XX
SQ   Sequence 594 BP; 132 A; 157 C; 155 G; 150 T; 0 U; 0 Other;
```

Alignment Scores:
```
Pred. No.:              2.4e-57      Length:          594
Score:                  750.00       Matches:         131
Percent Similarity:     99.24%       Conservative:    0
Best Local Similarity:  99.24%       Mismatches:      1
Query Match:            98.43%       Indels:          0
DB:                                  Gaps:            12
```

US-10-613-105-2 (1-132) x ADL57148 (1-594)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCT 78

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer 40
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 138

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy   61 LysArgCysCysGlyThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 259 CCAACCAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy 101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db 319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy 121 MetCysGlyLysSerCysValSerProValAla 132
        ||||||||||||||||||||||||||||||||||
Db 379 ATGTGTGGGAAATCCCGCGTTTCCCCTGTGAAAGCT 414

RESULT 33
ADK11941
ID   ADK11941 standard; DNA; 567 BP.
XX
AC   ADK11941;
XX
DT   06-MAY-2004 (first entry)
XX
DE   Breast cancer differentially expressed gene product #347.
XX
KW   ds; cytostatic; gene therapy; DKF2p566I133 activity inhibitor;
KW   breast cancer; differential expression.
XX
OS   Homo sapiens.
XX
PN   WO200305792 6-A1.
XX
PD   17-JUL-2003.
XX
PF   08-JAN-2003; 2003WO-US000657.
PR   08-JAN-2002; 2002US-0345637P.
XX
PA   (CHIR) CHIRON CORP.
XX
PI   Hansen R;
XX
DR   WPI; 2003-577534/54.
XX
PT   Inhibiting a cancerous phenotype of a cell, useful for treating breast
PT   cancer comprises contacting a cancerous mammalian cell with an agent for
PT   inhibition of DKF2p566I133 activity.
XX
PS   Claim 30; SEQ ID NO 347; 257pp; English.
XX
CC   The invention relates to a method of inhibiting a cancerous phenotype of
CC   a cell comprises contacting a cancerous mammalian cell with an agent for
CC   inhibition of DKF2p566I133 activity. The methods are useful for treating
CC   cancer, e.g. breast cancer. This sequence represents a gene product which
CC   is differentially expressed in breast cancer cells. The sequence can be
CC   used in the method of the invention.
XX
SQ   Sequence 567 BP; 127 A; 146 C; 150 G; 144 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:
Score: 6.29e-57
Percent Similarity: 745.00
Best Local Similarity: 100.00%
Query Match: 99.22%
DB: 97.77%

Length: 567
Matches: 128
Conservative: 1
Mismatches: 0
Indels: 0
Gaps: 0

US-10-613-105-2 (1-132) x ADK11941 (1-567)

Qy   4 SerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaProTrpAlaVal 23
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3 AGCGGCCTCTTCCCCTTCCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCTTGGGCTGTG 62

Qy   24 GluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSerAlaGlnCys 43
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63 GAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCTGCCCAGTGC 122

Qy   44 LeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCys 63
        ||||||||||||||||||||||||||||||||||||||||||||||:::||
Db 123 CTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAGAAGAAATGT 182

Qy   64 CysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArg 83
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 183 TGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAACCCAACCAGG 242

Qy   84 ArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsnProProAsn 103
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 243 AGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAACCCCCCCAAT 302

Qy 104 PheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMetCysGly 123
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 303 TTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGCATGTGTGGG 362

Qy 124 LysSerCysValSerProValLysAla 132
        |||||||||||||||||||||||||
Db 363 AAATCCTGCGTTTCCCCTGTGAAAGCT 389

RESULT 34
ADL57150
ID   ADL57150 standard; DNA; 594 BP.
XX
AC   ADL57150;
XX
DT   03-JUN-2004 (first entry)
XX
DE   Human NOV8f gene SEQ ID NO:95.
XX
KW   ds; gene; human; antidiabetic; anorectic; cardiant; hypotensive;
KW   antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
KW   protozoacide; nootropic; neuroprotective; antiparkinsonian;
KW   anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
KW   dermatological; antiasthmatic; antilipaemic; gene therapy;
KW   fibroblast growth factor receptor 4; FGFR4;
KW   complement factor I precursor; matrix metalloproteinase-15 precursor;
KW   MDC3; T-lymphocyte surface antigen Ly-9 precursor;
KW   fibroblast growth factor-21; FGF-21;
KW   alpha-2 macroglobulin-like polypeptide variant;
KW   antileukoproteinase 1 precursor; LIV-1; nuclear hormone receptor NOR-1;
KW   transmembrane protein-like; beta-neoendorphin-dynorphin precursor.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   19..417
FT   /*tag= a
XX
PN   WO2004022723-A2.
XX
PD   18-MAR-2004.
XX
PF   09-SEP-2003; 2003WO-US028141.

```
XX
PR   09-SEP-2002;  2002US-0409145P.
PR   10-SEP-2002;  2002US-0409544P.
PR   12-SEP-2002;  2002US-0410320P.
PR   16-SEP-2002;  2002US-0411060P.
PR   23-SEP-2002;  2002US-0412766P.
PR   23-SEP-2002;  2002US-0412825P.
PR   24-SEP-2002;  2002US-0412767P.
PR   25-SEP-2002;  2002US-0413342P.
PR   30-SEP-2002;  2002US-0414832P.
XX
PA   (CURA-)  CURAGEN CORP.
XX
PI   Zhong M,  Guo X,  Anderson DW,  Ort T,  Padigaru M,  Rieger DK;
XX
DR   WPI;  2004-315567/29.
DR   P-PSDB;  ADL57151.
XX
PT   New isolated NOVX polypeptides and polynucleotides, useful for
PT   preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT   osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT   asthma, or infections.
XX
PS   Claim 17; SEQ ID NO 95; 214pp; English.
XX
CC   The invention relates to a novel isolated polypeptide (NOVX) comprising a
CC   mature form of any of the 37 amino acid sequences fully defined in the
CC   specification. A polypeptide of the invention has antidiabetic,
CC   anorectic, cardiant, hypotensive, antiarteriosclerotic, anorectic,
CC   virucide, antibacterial, fungicide, protozoacide, nootropic,
CC   neuroprotective, antiparkinsonian, anticonvulsant, osteopathic,
CC   antiarthritic, antiinflammatory, dermatological, antiasthmatic, and
CC   antilipaemic activity. A polynucleotide of the invention may have a use
CC   in gene therapy. The polypeptides, nucleic acid molecules and antibodies
CC   are useful in the manufacture of a medicament for treating a syndrome
CC   associated with a human disease, preferably a NOVX-associated disorder.
CC   The nucleic acid molecules, polypeptides and antibodies are useful for
CC   treating, preventing or diagnosing diseases such as metabolic disorders,
CC   diabetes, obesity, infectious diseases (viral, bacterial, fungal,
CC   helminthic, and protozoal), anorexia, cancer, cardiovascular diseases
CC   (hypertension, atherosclerosis), neurodegenerative disorders, Alzheimer's
CC   disease, Parkinson's disease, epilepsy, immune disorders
CC   (osteoarthritis), haematopoietic disorders, inflammatory skin disorders,
CC   asthma, and various dyslipidaemias. The nucleic acids and polypeptides
CC   may also be used as targets for the identification of small molecules
CC   that modulate or inhibit e.g. neurogenesis, cell differentiation, cell
CC   proliferation, haematopoiesis, wound healing and angiogenesis, in gene
CC   therapy, in generation of antibodies that bind immunospecifically to NOVX
CC   substances for use in therapeutic or diagnostic methods. The nucleic
CC   acids are further used as hybridisation probes, in chromosome mapping,
CC   tissue typing, preventive medicine, and pharmacogenomics. The NOVX
CC   polypeptides of the invention show homology to certain known human
CC   proteins: NOV1a-1t show homology to fibroblast growth factor receptor 4
CC   (FGFR4); NOV2a shows homology to complement factor I precursor; NOV3a
CC   shows homology to matrix metalloproteinase-15 precursor; NOV4a shows
CC   homology to MDC3; NOV5a-5c show homology to T-lymphocyte surface antigen
CC   Ly-9 precursor; NOV6a-6m show homology to fibroblast growth factor-21
CC   (FGF-21); NOV7a-7c show homology to alpha-2 macroglobulin-like
CC   polypeptide variant; NOV8a-8g show homology to antileukoproteinase 1
CC   precursor; NOV9a-9i show homology to LIV-1 protein; NOV10a shows homology
CC   to nuclear hormone receptor NOR-1; NOV11a-11j show homology
CC   transmembrane protein-like; NOV12a-12c show homology to beta-neoendorphin
CC   -dynorphin precursor. The present sequence encodes a NOVX polypeptide of
CC   the invention.
XX
SQ   Sequence 594 BP; 132 A; 156 C; 155 G; 146 T; 0 U; 5 Other;

Alignment Scores:
Pred. No.:                               594
Score:                 5.86e-55          Matches:      127
Percent Similarity:    723.00            Conservative:   0
Best Local Similarity: 96.21%            Mismatches:     5
Query Match:           96.21%            Indels:         0
                       94.88%
```

```
DB:                      12          Gaps:        0

US-10-613-105-2 (1-112) x ADL57150 (1-594)

Qy    1  MetLysSerSerGlyLeuPheProPheValLeuLeuAlaLeuGlyThrLeuAlaPro   20
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19  ATGAAGTCCAGCGGCCTCTTCCCCTTCGTGCTGCTTGCCCTGGGAACTCTGGCACCT   78

Qy   21  TrpAlaValGluGlySerGlyLysPheLysAlaGlyValCysProProLysSer   40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79  TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT  138

Qy   41  AlaGlnCysLeuArgThrYrLysProGluCysGlnSerAspTrpGlnCysProGlyLys   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  139  GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG  198

Qy   61  LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn   80
            |||||||||||||||||||||||||||||||||||||||||||||||
Db  199  AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC  258

Qy   81  ProThrArgArgProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn  100
            ||||||||||||||||||||||||||||||||||||||||||||||
Db  259  CCAACAAGGAGGAAGCCTGGAAGTGCCCAGTGACTTATGGCCAATGTTYGATGCTTAAC  318

Qy  101  ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly  120
            ||||||||||||||||||||||||||||||||||||||||||||||
Db  319  CCCCCAAATTCTGAGATGGATGGCCAGTGCAAGCGTGACTGAAGTGTTGCATGGGC  378

Qy  121  MetCysGlyLysSerCysValSerProValLysAla  132
            ||||||||||||||||||||||||||||||||
Db  379  ATGTGTGGGAAATCCYGCGTTTCCCCTGTGAAAGCT  414
```

```
RESULT 35
AAL60894
ID   AAL60894 standard; cDNA; 605 BP.
XX
AC   AAL60894;
XX
DT   03-SEP-2003  (first entry)
XX
DE   Human secretory leukocyte proteinase inhibitor variant (Sepin) cDNA.
XX
KW   Human; forensic analysis; chromosome marker; organelle-specific marker;
KW   novel-related disorder; neurological disorder; gene therapy; nootropic;
KW   neuroprotective; secretory leukocyte proteinase inhibitor; inhibitor;
KW   variant; Sepin; gene; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..21
FT                   /*tag= a
FT   CDS             22..363
FT                   /*tag= b
FT                   /product= "Human Sepin protein"
FT   sig_peptide     22..96
FT                   /*tag= c
FT   mat_peptide     97..360
FT                   /*tag= d
FT                   /product= "Human mature Sepin protein"
FT   3'UTR           364..605
FT                   /*tag= e
FT   polyA_signal    565..570
FT                   /*tag= f
FT   polyA_site ,    590..605
FT                   /*tag= g
XX
PN   WO200304610-A2.
XX
PD   05-JUN-2003.
XX
PF   25-NOV-2002; 2002WO-EP013210.
```

XX
PR   28-NOV-2001; 2001US-0334147P.
PR   14-DEC-2001; 2001US-0340465P.
PR   18-APR-2002; 2002US-0373947P.
XX
PA   (GEST) GENSET SA.
XX
PI   Bejanin S, Tanaka H;
XX
DR   WPI; 2003-505202/47.
DR   P-PSDB; AAO30176.
XX
PT   New isolated polynucleotides and polypeptides useful as reagents in
PT   forensic analyses, as chromosome markers, as tissue/cell/organelle-
PT   specific markers, in producing expression vectors, or in screening and
PT   diagnostic assays.
XX
PS   Claim 1; Page 223-224; 242pp; English.
XX
CC   The invention relates to human novel polynucleotides and polypeptide
CC   sequences useful as reagents in forensic analyses, as chromosome markers,
CC   as tissue/cell/organelle-specific markers, in producing expression
CC   vectors, or in screening and diagnostic assays. The invention is used as
CC   reagents in screening and diagnostic assays for abnormal novel expression
CC   and/or biological activity, in screening compounds that may be used
CC   in the treatment of novel-related disorders, e.g. neurological disorders.
CC   The novel gene is also used in gene therapy. The present sequence is
CC   human secretory leukocyte proteinase inhibitor variant (Sepin) cDNA, a
CC   novel sequence of the invention
XX
SQ   Sequence 605 BP; 147 A; 154 C; 152 G; 152 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   605
Score: 9.96e-55   Length: 129
Percent Similarity: 97.73%   Matches: 0
Best Local Similarity: 97.73%   Conservative: 0
Query Match: 94.55%   Mismatches: 2
                                 Indels: 2
                                 Gaps: 1

US-10-613-105-2 (1-132) x AAL60894 (1-605)

Qy   1   MetLysSerSerGlyLeuPheProPheLeuValLeuAlaAlaLeuGlyThrLeuAlaPro   20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTCGCTGCCCTGGGAACTCTGGCACCT   81

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer   40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82 TGGGCTGTGGAAGGCGTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCTCCTCCTAAGAAATCT   141

Qy   41 AlaGlnCysLeuArgTyrLysGlyLysProGluCysProGlySerAspTrpGlnCysProGlyLys   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   142 GCCCAGTGCCTCCTTAGATACAAGGGAAAGCCTGAGTGCCCAGGTGACTGGCAGTGTCCAGGGAAG   201

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspProThrProAsn   80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   202 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC   261

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnLeuMetLeuAsn   100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   262 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC   321

Qy   101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   322 CCCCCAATTTCTGTGAGATGGATGGCCAGTGCGT-TAC---TTGAAGTGTTGCATGGGC   377

Qy   121 MetCysGlyLysSerCysValSerProValLysAla   132
          ||||||||||||||||||||||||||||||||||||
Db   378 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT   413

RESULT 36
ADK61172/c
ID   ADK61172 standard; DNA; 528 BP.

XX
AC   ADK61172;
XX
DT   06-MAY-2004   (first entry)
XX
DE   Ovarian cancer-related DNA #327 with altered ovarian cancer expression.
XX
KW   ds; gene; ovarian tumor; BRCA-1-like; BRCA-2-like; non-BRCA-like;
KW   gene expression; primer; cancer.
XX
OS   Homo sapiens.
XX
PN   WO2003068054-A2.
XX
PD   21-AUG-2003.
XX
PF   13-FEB-2003; 2003WO-US004688.
PR   13-FEB-2002; 2002US-0357031P.
XX
PA   (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA   (SLOK) SLOAN KETTERING INST CANCER RES.
XX
PI   Jazaeri AA, Boyd J, Liu ET;
XX
DR   WPI; 2003-689589/65.
XX
PT   Classifying an ovarian tumor as a BRCA-1-like or non-BRCA-
PT   like tumor by determining a pattern of expression in the ovarian tumor of
PT   several markers.
XX
PS   Disclosure; SEQ ID NO 342; 137pp; English.
XX
CC   The invention relates to a method of classifying an ovarian tumor as a
CC   BRCA-1-like or BRCA-2-like or non-BRCA-like tumor by: (1) determining a
CC   pattern of expression in the ovarian tumor of several markers given in
CC   the specification; and (2) comparing a similarity of the pattern of
CC   expression of the markers in the ovarian tumor to a pattern of expression
CC   of the markers in a comparison tissue of a known BRCA-1-like or BRCA-2-
CC   like or non-BRCA-like tumor. The method is useful for classifying an
CC   ovarian tumor as a BRCA-1-like or BRCA-2-like or non-BRCA-like tumor.
CC   This sequence corresponds to an ovarian cancer-related gene having an
CC   altered pattern of expression in ovarian cancer. (Note: The sequence data
CC   for this patent did not form part of the printed specification but was
CC   obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 528 BP; 130 A; 144 C; 130 G; 124 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   528
Score: 2.99e-53   Length: 120
Percent Similarity: 100.00%   Matches: 703.00
Best Local Similarity: 100.00% Conservative: 0
Query Match: 92.26%   Mismatches: 0
DB: Indels: 0
                                 Gaps: 0

US-10-613-105-2 (1-132) x ADK61172 (1-528)

Qy   13 LeuAlaLeuGlyThrLeuAlaProTrpAlaValGluGlyGlySerGlyLysSerPheLysLys   32
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   526 CTTGCCCTGGGAACTCTGGCACCTTGGGCTGTGGAAGGCGGCTCTGGAAAGTCCTTCAAAGCT   467

Qy   33 GlyValCysProProLysLysSerAlaGlnCysLeuArgTyrLysLysProGluCysGln   52
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   466 GGAGTCTGTCCTCCTAAGAAATCTGCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAG   407

Qy   53 SerAspTrpGlnCysProGlyLysLysArgCysCysValGlyIleLysLysCys   72
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   406 AGTGACTGGCAGTGTCCAGGGAAGAAGAGATGTTGTCCTGACACTTGTGGCATCAAATGC   347

Qy   73 LeuAspProValAspThrProAsnProThrArgArgLysProGlyLysCysProValThr   92
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Pred. No.:                              Length:        460
Score:                  3.41e-48        Matches:       112
Percent Similarity:     645.00          Conservative:    5
Best Local Similarity:  88.64%          Mismatches:     11
Query Match:            84.85%          Indels:          4
DB:                     84.65%          Gaps:            1
```

US-10-613-105-2 (1-132) x AAX16271 (1-460)

```
Qy    1 MetLysSerSerGlyLeuProPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
           |||       :::  :::  ||                              |||
Db   62 ATGAAAAGACAGCTATCGCGATCGCAGTGGCACCTGGTGGTTCGCACCGTAGCG---  118

Qy   21 TrpAlaValGluGlyGlySerPheLysAlaGlyValCysProProLysLysSer  40
              :::                    |||||||||||||||||||||||
Db  119 --------CAGGGCCTCTGGTAAAAGCTTCAAAGCTTGGCGATGCCGCGAAAAAATCC  169

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  170 GCGCAGTGTCTGCGGTACAAAAACCGAATGCCAGTCCGACTGGCAGTGCCCGGGTAAA  229

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysGlyAspAsnProThrProAsn  80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  230 AAACGTTGTTGCCCGGACACCTGCGGCATCAAATGCCTGGATCCGGTGATACCCCGAAC  289

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn  100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  290 CCGACTCGTCGAAAACCGGGTAAATGCCCGGTAACCTATGGCCAGTGTCTGATGCTGAAC  349

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  350 CCGCCGAACTTCTGCGAAATGGACGGCCAGTGTAAACGAGATCTGAAATGCTGTATGGGT  409

Qy  121 MetCysGlyLysLysSerCysValSerProValLysAla  132
         |||||||||||||||||||||||||||||||||||||||
Db  410 ATGTGCGGCAAATCTTGTTTTCCCCGTAAAAGCA  445
```

RESULT 38
AAC97622
ID   AAC97622 standard; DNA; 460 BP.
XX
AC   AAC97622;
XX
DT   27-FEB-2001 (first entry)
XX
DE   DNA encoding OmpA secretory leukocyte protease inhibitor.
XX
KW   Serine protease inhibitor; cytostatic; anti-inflammatory; arthritis;
KW   protease mediated tissue destruction; emphysema; glomerulonephritis;
KW   periodontitis; muscular dystrophy; tumour invasion; chymotrypsin;
KW   elastase; ds.
XX
OS   Synthetic.
XX
PN   US6132990-A.
XX
PD   17-OCT-2000.
XX
PF   07-JUN-1991;   91US-00712354.
XX
PR   06-DEC-1984;   84US-00678822.
PR   02-DEC-1985;   85US-00803471.
PR   29-JUL-1986;   86US-00890526.
PR   03-SEP-1986;   86US-00903471.
PR   30-MAR-1987;   87US-00031846.
PR   04-AUG-1987;   87US-00082962.
PR   03-JAN-1989;   89US-0293042.
XX
PA   (AMGE-) AMGEN BOULDER INC.
XX
PI   Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX
DR   WPI; 2000-678667/66.
XX
```
Db  346 CTGGATCCTGTTGACACCCCAAACCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACT  287
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   93 TyrGlyGlnCysLeuMetLeuAsnProProAsnPheCysGluMetAspGlyGlnCysLys  112
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  286 TATGGCCAAATGTTGATGCTTAACCCCCCAATTCTGTGAGATGGATGGCCAGTGCAAG  227

Qy  113 ArgAspLeuLysCysCysMetGlyMetCysGlyLysLysSerCysValSerProValLysAla  132
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  226 CGTGACTTGAAGTGTTGCATGGGCATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  167
```

RESULT 37
AAX16271
ID   AAX16271 standard; DNA; 460 BP.
XX
AC   AAX16271;
XX
DT   20-MAR-2003 (revised)
DT   19-APR-1999 (first entry)
XX
DE   DNA sequence of ompA-SLPI.
XX
KW   Serine protease inhibitor; human; emphysema; arthritis; peridontitis;
KW   muscular dystrophy; tumour invasion; glomerulonephritis; sepsis;
KW   acute leukemia; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   US5871956-A.
XX
PD   16-FEB-1999.
XX
PF   22-JUL-1994;   94US-00279056.
XX
PR   06-DEC-1984;   84US-00678822.
PR   29-JUL-1986;   86US-00890526.
PR   03-SEP-1986;   86US-00903471.
PR   30-MAR-1987;   87US-00031846.
PR   06-AUG-1990;   90US-00563832.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Stetler GL, Bandyopadhyay PK, Eisenberg SP, Thompson RC;
XX
DR   WPI; 1999-166640/14.
XX
PT   New DNA sequence encoding mammalian serine protease inhibitor - useful
PT   for producing inhibitors with different specificities,and
PT   treating diseases such as emphysema, arthritis, muscular dystrophy, and
PT   tumor invasion.
XX
PS   Example 3; Col 37-38; 37pp; English.
XX
CC   The present invention describes a DNA sequence (A) encoding an analogue
CC   of a mammalian serine protease inhibitor (B). The DNA sequences and
CC   recombinant methods allow manufacture of a class of inhibitors of e.g.
CC   cathepsin G, elastase, and trypsin, with different specificities. The
CC   recombinant serine protease product can be directed to act
CC   intracellularly or extracellularly and is useful in treating conditions
CC   caused by a disturbance in the native protease/protease inhibitor balance
CC   such as emphysema, arthritis, glomerulonephritis, peridontitis, muscular
CC   dystrophy, tumour invasion, sepsis, and acute leukemia. (A) allows (B) to
CC   be recombinantly produced in sufficient quantities and purities so as to
CC   provide economical pharmaceutical compositions. (B) is resistant to heat,
CC   acid, and a variety of proteolytic enzymes, is thermodynamically stable
CC   in extracellular conditions, and exhibits a high degree of self assembly
CC   forming an active tertiary structure in the absence of biochemical
CC   stimuli. The present sequence represents the DNA sequence of ompA-SLPI.
CC   (Updated on 20-MAR-2003 to correct PR field.)
XX
SQ   Sequence 460 BP; 123 A; 114 C; 125 G; 98 T; 0 U; 0 Other;

Alignment Scores:

PT   New serine protease inhibitors and DNA sequences for treating a protease-
PT   mediated condition or tissue destruction e.g. emphysema or tumor invasion
PT   and for recombinant production of inhibitors.
XX
PS   Example 3; Col 37-38; 47pp; English.
XX
CC   This invention relates to new purified and isolated mammalian serine
CC   protease inhibitor proteins which comprise at least 8 cysteine residues
CC   and no more than 107 amino acids. The protease inhibitors are capable of
CC   inhibiting chymotrypsin and elastase. Sequences AAC97526 - AAC97534,
CC   AAC97574 - AAC97581 and AAC97614 - AAC97641 represent DNA encoding the
CC   protease inhibitors of the invention, and include oligonucleotide
CC   sequences used in the isolation and characterisation of the proteins.
CC   Primers AAC97535 - AAC97573 and AAC95582 - AAC97613 are used in the
CC   construction of DNA encoding the protease inhibitors. Peptide sequences
CC   AAB53098 - AAB53122 represent the protease inhibitor of the invention and
CC   various other peptides used in the isolation of the protease inhibitors.
CC   The serine protease inhibitors have cytostatic and anti-inflammatory activity.
CC   The serine protease inhibitor protein is useful for treating a protease-
CC   mediated condition, which includes protease mediated tissue destruction,
CC   e.g. emphysema, arthritis, glomerulonephritis, periodontitis, muscular
CC   dystrophy or tumour invasion. It is also useful for modulating protease
CC   activity. The DNAs are useful for producing the proteins, especially by
CC   recombinant methods
XX
SQ   Sequence 460 BP; 123 A; 114 C; 125 G; 98 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   3.41e-48
Score: 645.00
Percent Similarity: 88.64%   Length: 460
Best Local Similarity: 84.85%   Matches: 112
Query Match: 84.65%   Conservative: 5
DB: Mismatches: 11
                                        Indels: 4
                                        Gaps: 1

US-10-613-105-2 (1-132) x AAC97622 (1-460)

Qy   1   MetLysSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
          ||||||| |||| :: ||| ||||
Db   62 ATGAAAAGACAGCTATCGCGATCGCAGTGGCACTGGCACTGGTTCGCTACCGTAGCG--- 118

Qy   21 TrpAlaValGluGlyLysSerPheLysAlaGlyValCysProProLysSerSer 40
                    |||| |||||||| ||||||||
Db   119 -------CAGGCCTCGTAAAAGCTTCAAAGTGGCGTATGCCCGCCGAAAAAATCC 169

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
          ||||||| ::: ::: ||| ::: |||
Db   170 GCGCAGTGTCTGCGGTACAAAAACCGGGTAAATGCCCGACTGGCAGTGCCCGGGTAAA 229

Qy   61 LysArgCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
          |||||||||||||| ||||||
Db   230 AAACGTTGTCCGGACACCTGCGGCATCAAATGCCTGGATCCGGTTGATACCCCGAAC 289

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
          ||||||| |||||| |||||||||||||
Db   290 CCGACTCGTCGAAAACCGGGTAAATGCCCGGTAACCTATGGCCAGTGTCTGATGCTGAAC 349

Qy   101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
          |||||||||||| |||||| ||||||
Db   350 CCGCCGAACTTCTGCGAAATGGACGGCCAGTGTAAACGCGATCTGAAATGCTGTATGGGT 409

Qy   121 MetCysGlyLysSerCysValSerProValLysAla 132
          |||||||||||||
Db   410 ATGTGCGGCAAATCTTGTGTTTCCCCGGTAAAAGCA 445

RESULT 39
AAI67582
ID   AAI67582 standard; DNA; 460 BP.
XX
AC   AAI67582;
XX
DT   11-FEB-2002 (first entry)
XX

DE   DNA sequence coding for ompA SLPI.
XX
KW   Serine protease inhibitor protein; recombinant; leukocyte elastase;
KW   trypsin; secretory leukocyte protease inhibitor; SLPI; ds.
XX
OS   Homo sapiens.
XX
PN   US6291662-B1.
XX
PD   18-SEP-2001.
XX
PF   22-SEP-1998;   98US-0015808085.
XX
PR   05-DEC-1984;   84US-00678222.
PR   02-DEC-1985;   85US-00803471.
PR   29-JUL-1986;   86US-00890526.
PR   30-MAR-1987;   87US-00011846.
PR   06-AUG-1990;   90US-00563832.
PR   22-JUL-1994;   94US-00279056.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Bandyopadhyay PK, Eisenberg SP, Stetler GL, Thompson RC;
XX
DR   WPI; 2001-63797473.
XX
PT   New DNA sequences, useful in recombinant DNA techniques for directing the
PT   production of a serine protease inhibitor protein, e.g. leukocyte
PT   elastase or trypsin.
XX
PS   Example 3; Col 36; 37pp; English.
XX
CC   The invention relates to a DNA sequence encoding an analog of a mammalian
CC   serine protease inhibitor protein. The analog comprises at least eight
CC   cysteine residues and possesses serine protease inhibitor activity, and
CC   at least one operational element consisting of a promoter, an operator, a
CC   leader sequence, a Shine-Dalgarno sequence, a ribosome binding site or a
CC   terminator codon. The DNA is useful in recombinant DNA techniques for
CC   directing the production of a serine protease inhibitor protein, e.g.
CC   leukocyte elastase or trypsin. The present sequence represents the DNA
CC   sequence coding for ompA-secretory leukocyte protease inhibitor (SLPI)
XX
SQ   Sequence 460 BP; 123 A; 114 C; 125 G; 98 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   3.41e-48
Score: 645.00
Percent Similarity: 88.64%   Length: 460
Best Local Similarity: 84.85%   Matches: 112
Query Match: 84.65%   Conservative: 5
DB: Mismatches: 11
                                        Indels: 4
                                        Gaps: 1

US-10-613-105-2 (1-132) x AAI67582 (1-460)

Qy   1   MetLysSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
          ||||||| |||| :: ||| ||||
Db   62 ATGAAAAGACAGCTATCGCGATCGCAGTGGCACTGGCACTGGTTCGCTACCGTAGCG--- 118

Qy   21 TrpAlaValGluGlyLysSerPheLysAlaGlyValCysProProLysSer 40
                    |||| |||||||| ||||||||
Db   119 -------CAGGCCTCGTAAAAGCTTCAAAGTGGCGTATGCCCGCCGAAAAAATCC 169

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
          ||||||| ::: ::: ||| ::: |||
Db   170 GCGCAGTGTCTGCGGTACAAAAACCGGGTAAATGCCCGACTGGCAGTGCCCGGGTAAA 229

Qy   61 LysArgCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
          |||||||||||||| ||||||
Db   230 AAACGTTGTCCGGACACCTGCGGCATCAAATGCCTGGATCCGGTTGATACCCCGAAC 289

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
          ||||||| |||||| |||||||||||||
Db   290 CCGACTCGTCGAAAACCGGGTAAATGCCCGGTAACCTATGGCCAGTGTCTGATGCTGAAC 349

```
Qy    101 ProProAsnPheCysGluMetAspGlyGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    350 CCGCCGAACTTCTGCGAAATGGACGGCCAGTGTAAACGAGATCTGAAATGCTGTATGGGT 409

Qy    121 MetCysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||||||||
Db    410 ATGTGCGGCAAATCCTGTGTTCCCCGGTAAAAGCA 445
```

RESULT 40
ABK88025
ID   ABK88025 standard; DNA; 1525 BP.
XX
AC   ABK88025;
XX
DT   07-OCT-2002 (first entry)
XX
DE   DNA sequence encoding rSLAP1 fusion protein.
XX
KW   rSLAP1; gene; ds; Alzheimer's disease; tumour angiogenesis; malaria;
KW   emphysema; asthma; chronic obstructive pulmonary disease;
KW   cystic fibrosis; otitis media; otitis external; HIV; psoriasis; eczema;
KW   human immunodeficiency virus; atopic dermatitis; muscular dystrophy;
KW   herpes; ulceration; sepsis; rheumatoid arthritis; periodontal disease;
KW   tumour metastasis; osteoporosis; Paget's disease; scleroderma;
KW   glomerulonephritis; hypertension.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   RBS   6..8
FT   /tag= a
FT   /standard_name= "Ribosome binding site"
FT   CDS   9..1520
FT   /tag= b
FT   /product= "rSLAP1 fusion protein"
FT   misc_feature   12..1193
FT   /tag= c
FT   /note= "AAT coding region"
FT   misc_feature   1194..1196
FT   /tag= d
FT   /note= "linking codon"
FT   misc_feature   1197..1517
FT   /tag= e
FT   /note= "SLPI coding region"
XX
PN   WO200250287-A2.
XX
PD   27-JUN-2002.
XX
PF   18-DEC-2001; 2001WO-US049256.
XX
PR   18-DEC-2000; 2000US-0256699P.
PR   20-NOV-2001; 2001US-0331966P.
XX
PA   (ARRI-) ARRIVA PHARM INC.
XX
PI   Barr PJ, Gibson HL, Pemberton P;
XX
DR   WPI; 2002-500631/53.
DR   P-PSDB; AAU99884.
XX
PF   Example 3; Page 89-90; 134pp; English.
XX
PT   Novel fusion protein useful for inhibiting protease activity associated
PT   with a disorder such as emphysema, asthma, comprises a first protease
PT   inhibitor comprising alpha 1-antitrypsin and a second protease inhibitor.
XX
CC   This invention relates to a novel fusion protein comprising a first
CC   protease inhibitor comprising an alpha1-antitrypsin or its functionally
CC   active portion and a second protease inhibitor or its functionally active
CC   protein. The fusion proteins of the invention may act as an inhibitor of

CC   protease activity. The fusion protein of the invention is useful for
CC   inhibiting protease activity associated with a disorder such as
CC   emphysema, asthma, chronic obstructive pulmonary disease, cystic
CC   fibrosis, otitis media, otitis external or HIV infection, or for treating
CC   an individual suffering from or at risk for a disease or disorder
CC   involving unwanted protease activity. The proteins are useful for
CC   treating dermatological diseases such as atopic dermatitis, eczema and
CC   psoriasis, in inflammatory responses to viral infection, and for treating
CC   herpes infection, corneal or epidermal ulceration, chronic non-healing
CC   wounds, sepsis, rheumatoid arthritis, periodontal disease, tumour
CC   metastasis and tumour angiogenesis, gastric ulceration, osteoporosis,
CC   Paget's disease, glomerulonephritis, scleroderma, malaria, bacterial
CC   infection, Alzheimer's disease, hypertension and muscular dystrophy. The
CC   present sequence represents the DNA encoding the rSLAP1 fusion protein of
CC   the invention
XX
SQ   Sequence 1525 BP; 467 A; 287 C; 314 G; 457 T; 0 U; 0 Other;
```

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 5.3e-47 | Length: | 1525 |
| Score: | 638.50 | Matches: | 112 |
| Percent Similarity: | 87.97% | Conservative: | 5 |
| Best Local Similarity: | 84.21% | Mismatches: | 9 |
| Query Match: | 83.79% | Indels: | 7 |
| DB: | | Gaps: | 1 |

```
US-10-613-105-2 (1-132) x ABK88025 (1-1525)

Qy      7 PheProPheLeuValLeu------------------LeuAlaLeuGlyThrLeuAla  19
          |||  |||||||||||:::::::                       |||   :::
Db   1119 TTCGTTTTCTGATGATCGAGCAGACACTAAAGCCCATTGTTTATGGGTAAGGTTGTC 1178

Qy     20 ProTrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLys  39
                                          :::         ||||||||
Db   1179 AACCCAACTCAGAAGATGTCCGGAAAGTCTTCAAGGCCGGTGTTTGTCCACCAAAGAAG 1238

Qy     40 SerAlaGlnCysLeuArgTyrLysProGluCysGlnSerAspTrpGlnCysProGly  59
          |||||                    |||   |||            ||||||||||
Db   1239 TCCGCTCAATGTTGAGATACAAGAACCCAGAATGTCAATCCGACTGGCAATGTCCAGT 1298

Qy     60 LysLysArgCysCysProAspThrCysGlyIleLeuLysAspProValAspThrPro  79
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1299 AAGAAGAGATGTTGTCCAGACACTTGTGGTATCAAGTGTCAGACCCCAGTTGACACCCCA 1358

Qy     80 AsnProThrArgArgLysProGlyLysCysProValThrTyrTyrGlnCysLeuMetLeu  99
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1359 AACCCAACTAGAAGAAAAGCCAGGTAAGTGTCCAGTTACTTACGGTCAATGTTGATGGTG 1418

Qy    100 AsnProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMet 119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1419 AACCCACCAAACTTCTGTGAAATGGACGGTCAATGTAAGAGAGACTTGAAGTGTTGTATG 1478

Qy    120 GlyMetCysGlyLysSerCysValSerProValLysAla 132
          |||||||||||||||||||||||||||||||||||||
Db   1479 GGTATGTGTGGTAAGTCCTGTGTTCCCCAGTCAAGGCC 1517

RESULT 41
ABK88016
ID    ABK88016 standard; DNA; 321 BP.
XX
AC    ABK88016;
XX
DT    07-OCT-2002 (first entry)
XX
DE    DNA encoding human secretory leukoprotease inhibitor (SLPI) protein.
XX
KW    secretory leukoprotease inhibitor; human; gene; ds; protease inhibitor;
KW    malaria; emphysema; asthma, chronic obstructive pulmonary disease; SLPI;
KW    cystic fibrosis; otitis media; otitis external; HIV; psoriasis; eczema;
KW    human immunodeficiency virus; atopic dermatitis; muscular dystrophy;
KW    herpes; ulceration; sepsis; rheumatoid arthritis; periodontal disease;
KW    tumour metastasis; tumour angiogenesis; osteoporosis; Paget's disease;
KW    glomerulonephritis; scleroderma; Alzheimer's disease; hypertension.
```

```
                                                           Qy    66 AspThrCysGlyIleLysCysLeuAspProThrArgArgLys   85
                                                                    |||||||||||||||||||||||||||||||||||||||||||
                                                           Db   121 GACACTTGTGGTATCAAGTGTCTAGACCCAGTTGACCAGAAGAAAG  180

                                                           Qy    86 ProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsnProProAsnPheCys  105
                                                                    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                                                           Db   181 CCAGGTAAGTGCCCAGTTACTTACGGTCAATGTTTGATGTTGAACCCACCAAACTTCTGT  240

                                                           Qy   106 GluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMetCysGlyLysSer  125
                                                                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                                                           Db   241 GAAATGGACGGTCAATGTAAGAGAGACTTGAAGTGTTGTATGGGTATGTGTGGTAAGTCC  300

                                                           Qy   126 CysValSerProValLysAla  132
                                                                    |||||||||||||||||||||
                                                           Db   301 TGTGTTTCCCCAGTCAAGGCC  321
```

Left column (first entry):

```
XX   WPI; 2003-513974/48.
DR   P-PSDB; ADE95614.
XX
PT   New NOVX polypeptides and nucleic acids, useful for preventing or
PT   treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT   atherosclerosis or diabetes, and in chromosome mapping, tissue typing or
PT   pharmacogenomics.
XX
PS   Claim 20; SEQ ID NO 145; 211pp; English.
XX
CC   This invention relates to novel NOVX proteins, and the DNA sequence which
CC   encode them, having properties related to stimulation of biochemical or
CC   physiological responses in a cell, a tissue, an organ or an organism.
CC   Compounds which modulate the proteins of the invention may have cardiant,
CC   antiarteriosclerotic, hypotensive, cytostatic, anorectic, antirheumatic,
CC   antiarthritic, antidiabetic, nephrotropic, dermatological,
CC   immunosuppressive, anti-HIV, antiinflammatory, neuroprotective,
CC   nootropic, antipsoriatic, antiparkinsonian, antiasthmatic, neuroleptic,
CC   antidepressant, antiallergic or gynaecological activities. The DNA
CC   sequences of the invention may be useful for gene therapy whilst the
CC   protein sequences may allow the development of a vaccine. The protein is
CC   useful in the manufacture of a medicament for treating a syndrome
CC   associated with a human disease. The invention may be useful in
CC   diagnosing, treating or preventing NOVX-associated disorders, for example
CC   cardiomyopathy, atherosclerosis, hypertension, cancer, obesity,
CC   rheumatoid arthritis, diabetes, glomerulonephritis, psoriasis, skin
CC   disorders, AIDS, inflammation, multiple sclerosis, Alzheimer's disease,
CC   Parkinson's disease, asthma, schizophrenia, depression, allergies or
CC   fertility disorders. The nucleic acids may further be used as
CC   hybridisation probes, in chromosome mapping, tissue typing, preventive
CC   medicine, and pharmacogenomics. The present sequence is the cDNA sequence
CC   which encodes the human NOVX27c protein of the invention.
XX
SQ   Sequence 321 BP; 81 A; 76 C; 87 G; 77 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                        Length:           321
Score:            1.71e-47        Matches:          107
Percent Similarity:  100.00%      Conservative:     0
Best Local Similarity: 100.00%    Mismatches:       0
Query Match:      83.33%          Indels:           0
DB:               10              Gaps:             0

US-10-613-105-2 (1-132) x ADE95613 (1-321)

Qy    26  SerGlyLysSerPheLysAlaGlyValCysProProLysLysSerAlaGlnCysLeuArg  45
Db     1  TCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCTGCCCAGTGCCTTAGA  60

Qy    46  TyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysPro  65
Db    61  TACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAGAAGAGATGTTGTCCT  120

Qy    66  AspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArgArgLys  85
Db   121  GACACCTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAACCCAACAAGGAGGAAG  180

Qy    86  ProGlyLysCysProValThrTyrGlnCysLeuMetLeuAsnProProAsnPheCys  105
Db   181  CCTGGGAAGTGCCCGGTAACCTATCAGTGCCTAATGCTTAACCCCCCAAATTTCTGT  240

Qy   106  GluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMetCysGlyLysSer  125
Db   241  GAGATGGATGGGCAAGCTGCAAGTGTTGCATGGGCATGTGTGGGAAATCC  300

Qy   126  CysValSerProValLysAla  132
Db   301  TGCGTTTCCCCTGTGAAAGCT  321

RESULT 43
AAN60463
ID   AAN60463 standard; DNA; 324 BP.
```

Right column (second entry):

```
XX   AAN60463;
AC
XX
DT   25-MAR-2003  (revised)
DT   01-JAN-1980  (first entry)
XX
DE   Serine protease inhibitor analogue having similar properties to a protein
DE   isolated from parotid secretions.
XX
KW   Serine protease-inhibitor; parotid; leukocyte-elastase-inhibitor;
KW   trypsin-; inhibitor; 88.
XX
OS   Synthetic.
XX
PN   WO8603519-A.
XX
PD   19-JUN-1986.
XX
PF   04-DEC-1985;    85WO-US002385.
XX
PR   06-DEC-1984;    84US-00678822.
PR   02-DEC-1985;    85US-00803471.
XX
PA   (SYND ) SYNERGEN BIOLOGICAL INC.
XX
PI   Bandyopadh PK, Eisenberg SP, Stetler GL, Thompson RC;
XX
DR   WPI; 1986-169458/26.
DR   P-PSDB; AAP60562.
XX
PT   New synthetic DNA sequences for directing microbial synthesis - for
PT   prodn. of single poly:peptide chain serine protease inhibitor having
PT   leukocyte elastase and trypsin inhibitory sites.
XX
PS   Disclosure; Page 14; 59pp; English.
XX
CC   The sequence directs synthesis of a single chain polypeptide serine
CC   protease-inhibitor, which believed to have at least 2 active sites, 1
CC   exhibiting leukocyte-elastase-inhibiting properties and the other
CC   exhibiting activity against trypsin. See also AAN60464-69 and AAP60562-
CC   66. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 324 BP; 87 A; 85 C; 88 G; 64 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                        Length:           324
Score:            1.73e-47        Matches:          107
Percent Similarity:  100.00%      Conservative:     0
Best Local Similarity: 100.00%    Mismatches:       0
Query Match:      83.33%          Indels:           0
DB:               1               Gaps:             0

US-10-613-105-2 (1-132) x AAN60463 (1-324)

Qy    26  SerGlyLysSerPheLysAlaGlyValCysProProLysLysSerAlaGlnCysLeuArg  45
Db     1  AGCGGTAAAAGCTTCAAAGCTGGGTATGCGCCGCCGAAAAAATCCGGCCAGTGTCTGCGG  60

Qy    46  TyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysPro  65
Db    61  TACAAAAAACCGGAATGCCAGTCCGACTGGCAGTGCCCGGGTAAAAAACGTTGTTGCCCG  120

Qy    66  AspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArgArgLys  85
Db   121  GACACCTGCGGCATCAAATGCCTGGATCCGGTTGATACCCCGAACCCGACCCGTCGAAAA  180

Qy    86  ProGlyLysCysProValThrTyrGlnCysLeuMetLeuAsnProProAsnPheCys  105
Db   181  CCGGGTAAATGCCCGGTAACCTATCAGTGTCTGATGCTGAATCCGCCGAACTTCTGC  240

Qy   106  GluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMetCysGlyLysSer  125
Db   241  GAAATGGACGGCCAGTGTAAACGAGATCTGAAATGCTGTATGGGTATGTGCGGCAAATCT  300
```

```
Qy          126 CysValSerProValLysAla 132
                |||||||||||||||||||||
Db          301 TGTGTTTCCCCGGTAAAAGCA 321

RESULT 44
AAN60464
ID   AAN60464 standard; DNA; 324 BP.
XX
AC   AAN60464;
XX
DT   25-MAR-2003  (revised)
DT   01-JAN-1980  (first entry)
XX
DE   Synthetic sequence capable of directing microbial synthesis of a
DE   secretory leukocyte protease-inhibitor.
XX
KW   Serine protease-inhibitor; parotid; leukocyte-elastase-inhibitor;
KW   trypsin-; inhibitor; ss.
XX
OS   Synthetic.
XX
PN   WO8603519-A.
XX
PD   19-JUN-1986.
XX
PF   04-DEC-1985;                85WO-US002385.
XX
PR   06-DEC-1984;                84US-00678822.
PR   02-DEC-1985;                85US-00803471.
XX
PA   (SYND ) SYNERGEN BIOLOGICAL INC.
XX
PI   Bandyopadh PK,  Eisenberg SP,  Stetler GL,  Thompson RC;
XX
DR   WPI; 1986-169458/26.
DR   P-PSDB; AAP60563.
XX
PT   New synthetic DNA sequences for directing microbial synthesis - for
PT   prodn. of single poly:peptide chain serine protease inhibitor having
PT   leukocyte elastase and trypsin inhibitory sites.
XX
PS   Disclosure: Page 15; 59pp; English.
XX
CC   The sequence directs synthesis of a secretory leukocyte protease-
CC   inhibitor. See also AAN60463, AAN60465-69 and AAP60562. AAP60564-66.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 U; 0 Other;

Alignment Scores:   -
Pred. No.:                        Length:              324
Score:             1.73e-47       Matches:             107
                   635.00         Conservative:          0
Percent Similarity: 100.00%       Mismatches:            0
Best Local Similarity: 100.00%    Indels:                0
Query Match:        83.33%        Gaps:                  0
DB:                 1

US-10-613-105-2 (1-132) x AAN60464 (1-324)

Qy   26 SerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeuArg 45
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TCTGGTAAAAGCTTCAAAGCTGGCGTATGCCCGCCGAAAAAATCCGCGCAGTGTCTGCGG 60

Qy   46 TyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysPro 65
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TACAAAAAACCCGAATGCCAGTCCGACTGGCAGTGCCCGGGTAAAAAACGTTGTTGCCCG 120

Qy   66 AspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArgArgLys 85
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GACACCTGCGGCATCAAATGCCTGGATCCGGTTGATACCCCGAACCCGACTCGTCGAAAA 180

Qy   86 ProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsnProProAsnPheCys 105
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      181 CCGGGTAAATGCCCGGTAACCTATGGCCAGTGTCTGATGCTGAACCCGCCGAACTTCTGC 240
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      106 GluMetAspGlyGlnCysGlyLysLysArgAspLeuLysCysCysMetGlyMetCysGlyLysGlyLysSer 125
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 GAAATGACGGCCAGTGAAACGAGATCTGAAATGCTGTATGTGTGCGGCAAATCT 300
```

```
Qy      126 CysValSerProValLysAla 132
                |||||||||||||||||||||
Db      301 TGTGTTTCCCCGGTAAAAGCA 321

RESULT 45
AAX16236
ID   AAX16236 standard; DNA; 324 BP.
XX
AC   AAX16236;
XX
DT   20-MAR-2003  (revised)
DT   19-APR-1999  (first entry)
XX
DE   Serine protease inhibitor direct manufacturing DNA sequence (SLPI).
XX
KW   Serine protease inhibitor; human; emphysema; arthritis; peridontitis;
KW   muscular dystrophy; tumour invasion; glomerulonephritis; sepsis;
KW   acute leukemia; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   US5871956-A.
XX
PD   16-FEB-1999.
XX
PF   22-JUL-1994;                94US-00279056.
XX
PR   06-DEC-1984;                84US-00678822.
PR   29-JUL-1986;                86US-00890526.
PR   03-SEP-1986;                86US-00903471.
PR   30-MAR-1987;                87US-00031846.
PR   06-AUG-1990;                90US-00563832.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Stetler GL,  Bandyopadhyay PK,  Eisenberg SP,   Thompson RC;
XX
DR   WPI; 1999-166640/14.
XX
PT   New DNA sequence encoding mammalian serine protease inhibitor - useful
PT   for recombinantly producing inhibitors with different specificities and
PT   treating diseases such as emphysema, arthritis, muscular dystrophy, and
PT   tumor invasion.
XX
PS   Example 2; Col 29; 37pp; English.
XX
CC   The present invention describes a DNA sequence (A) encoding an analogue
CC   of a mammalian serine protease inhibitor (B). The DNA sequences and
CC   recombinant methods allow manufacture of a class of inhibitors of e.g.
CC   cathepsin G, elastase, and trypsin, with different specificities. The
CC   recombinant serine protease product can be directed to act
CC   intracellularly or extracellularly and is useful in treating conditions
CC   caused by a disturbance in the native protease/protease inhibitor balance
CC   such as emphysema, arthritis, glomerulonephritis, peridontitis, muscular
CC   dystrophy, tumour invasion, sepsis, and acute leukemia. (A) allows (B) to
CC   be recombinantly produced in sufficient quantities and purities so as to
CC   provide economical pharmaceutical compositions. (B) is resistant to heat,
CC   acid, and a variety of proteolytic enzymes, is thermodynamically stable
CC   in extracellular conditions, and exhibits a high degree of self assembly
CC   forming an active tertiary structure in the absence of biochemical
CC   stimuli. The present sequence represents a DNA sequence which is used to
CC   direct manufacture of recombinant serine protease inhibitors. (Updated on
CC   20-MAR-2003 to correct PR field.)
XX
SQ   Sequence 324 BP; 86 A; 85 C; 87 G; 66 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:              1.73e-47            Length:         324
Score:                  635.00              Matches:        107
Percent Similarity:     100.00%             Conservative:   0
Best Local Similarity:  100.00%             Mismatches:     0
Query Match:            83.33%              Indels:         0
DB:                     2                   Gaps:           0

US-10-613-105-2 (1-132) x AAX16236 (1-324)

Qy    26 SerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeuArg 45
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TCTGGTAAAAGCTTCAAAGCTGGGCTATGCCCGCCGAAAAAATCCGCGCAGTGTCTGCGG 60

Qy    46 TyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysPro 65
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TACAAAAAACCGGAATGCCAGTCCGACTGGCAGTGCCGGGTAAAAAAACGTTGTTGCCCG 120

Qy    66 AspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArgArgLys 85
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GACACCTGCGGCATCAAATGCCTGGATCCGGTTGATACCCCGAACCCGACTCGTCGAAAA 180

Qy    86 ProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsnProProAsnPheCys 105
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCGGGTAAATGCCCGGTAACCTATGGCCAGTGTCTGATGCTGAACCCGCCGAACTTCTGC 240

Qy   106 GluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMetCysGlyLysSer 125
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GAAATGGACGGCCAGTGTAAACGTGATCTGAAATGCTGTATGGGTATGTGCGGCAAATCT 300

Qy   126 CysValSerProValLysAla 132
            ||||||||||||||||||||
Db   301 TGTGTTTCCCCGGTAAAAGCA 321
```

Search completed: October 24, 2004, 04:07:34
Job time : 450 secs

This Page Blank (1924)

OM protein - nucleic search, using frame_plus_p2n model

```
          GenCore version 5.1.6
      Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:    October 24, 2004, 03:15:16 ; Search time 3545 Seconds
                                        (without alignments)
                                        1760.858 Million cell updates/sec

Title:          US-10-613-105-2
Perfect score:  762
Sequence:     1 MKSSGLFPFLVLLALGTLAP.........RDLKCCMGMCGKSCVSPVKA 132

Scoring table:  BLOSUM62

Xgapop  10.0 ,  Xgapext  0.5
Ygapop  10.0 ,  Ygapext  0.5
Fgapop   6.0 ,  Fgapext  7.0
Delop    6.0 ,  Delext   7.0

Searched:     4526729 seqs, 2364484745 residues

Total number of hits satisfying chosen parameters:     9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 100 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US10613105/runat_22102004_170440_18687/app_query.fasta_1.327
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=45 -MODE=LOCAL
-OUTFMT=pto -NORM=6xt -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10613105 @CGN_1_1_3731 @runat_22102004_170440_18687 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=120 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -XGAPEXT=0.5 -FGAPOP=6 -DELOP=6 -DELEXT=7

Database :    GenEmbl:*
              1: gb_ba:*
              2: gb_htg:*
              3: gb_in:*
              4: gb_om:*
              5: gb_ov:*
              6: gb_pat:*
              7: gb_ph:*
              8: gb_pl:*
              9: gb_pr:*
             10: gb_ro:*
             11: gb_sts:*
             12: gb_sy:*
             13: gb_un:*
             14: gb_vi:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	762	100.0	399	6	AX302535	AX302535 Sequence
2	762	100.0	399	9	AF114471	AF114471 Homo sapi
3	762	100.0	551	6	AX778065	AX778065 Sequence
4	762	100.0	565	6	I01502	I01502 Sequence 3
5	762	100.0	565	6	AX069252	AX069252 Sequence
6	762	100.0	573	9	HSSLIPR	X04503 Human SLPI
7	762	100.0	594	6	CQ727371	CQ727371 Sequence
8	762	100.0	594	6	AR270708	AR270708 Sequence
9	762	100.0	594	6	AX328399	AX328399 Sequence
10	762	100.0	594	6	AX334507	AX334507 Sequence
11	762	100.0	594	6	AX335376	AX335376 Sequence
12	762	100.0	594	6	AX577960	AX577960 Sequence
13	762	100.0	599	6	AX780010	AX780010 Sequence
14	762	100.0	625	9	HSALPR	X04470 Human mRNA
15	762	100.0	1084	9	BD203743	BD203743 Human nuc
16	762	100.0	605	6	AX014898	AX014898 Sequence
17	720.5	94.6	1525	9	BC020708	BC020708 Homo sapi
18	638.5	83.8	321	6	CQ413516	CQ413516 Sequence
19	635	83.3	324	6	AX772818	AX772818 Sequence
20	635	83.3	1525	6	AX670655	AX670655 Sequence
21	635	83.3	321	6	AX670643	AX670643 Sequence
22	632	82.9	321	6	I08404	I08404 Sequence 1
23	623	81.8	321	6	AX670647	AX670647 Sequence
24	623	81.8	321	6	A08111	A08111 Synthetic A
25	623	81.8	321	6	A08106	A08106 Synthetic A
26	621	81.5	614	6	A08107	A08107 Synthetic A
27	621	81.5	614	6	A08112	A08112 Synthetic A
28	621	81.5	321	6	A08108	A08108 Synthetic A
29	614	80.6	2657	6	A08110	A08110 Synthetic A
30	614	80.6	498	6	A08105	A08105 Synthetic A
31	547	71.8	63796	6	A08109	A08109 Synthetic A
32	545	71.5	600	9	HSSLIPG	X04502 Human SLPI
33	537.5	70.5	490	6	I01501	I01501 Sequence 2
34 c	508.5	66.7	667	9	HS300I2	AL035660 Human DNA
35	482.5	63.3	680	4	PIGALP	M57446 Porcine ant
36	471	61.8	396	10	AF178426	AF178426 Rattus no
37	471	61.8	396	10	AF151982	AF151982 Rattus no
38	463.5	60.8	409	10	AF421377	AF421377 Rattus no
39	463.5	60.8	671	6	AX642233	AX642233 Sequence
40 c	463.5	60.8	682	6	AX642236	AX642236 Sequence
41	463.5	60.8	894	6	AX577885	AX577885 Sequence
42	463.5	60.8	1123	10	MMU88093	U88093 Mus musculu
43	463.5	60.8	325	10	MMU94341	U94341 Mus musculu
44	463.5	60.8	194	10	BC028509	BC028509 Mus muscu
45	458	60.1	2737	10	MMU73004	U73004 Mus musculu
46	348	45.7	281747	6	AR270461	AR270461 Sequence
47	316.5	41.5	296417	6	AR067991	AR067991 Sequence
48	302	39.6	4090	4	AY346135	AY346135 Ovis arie
49 c	302	39.6	313	10	AC094865	AC094865 Rattus no
50 c	285	37.4	313	10	AC112730	AC112730 Rattus no
51	283.5	37.2	2435	10	AF205374	AF205374 Mus muscu
52	283.5	37.2	190669	6	CQ395755	CQ395755 Sequence
53	277.5	36.4	212105	6	CQ402085	CQ402085 Sequence
54	277.5	36.4	411	10	AF002719	AF002719 Mus muscu
55	265	34.8	411	10	AL590429	AL590429 Mouse DNA
56 c	264	34.6	411	10	AC132741	AC132741 Rattus no
57 c	264	34.6	222	6	AX334505	AX334505 Sequence
58 c	264	34.6	188	6	AX335289	AX335289 Sequence
59 c	261	34.3	292	6	AX335639	AX335639 Sequence
60 c	243	31.9	292	4	PIGUF52	M31216 Porcine ute
61	224	29.4	292	6	CQ408469	CQ408469 Sequence
62	224	29.4	292	6	BD264991	BD264991 Compositi
63	224	29.4	292	6	AR238394	AR238394 Sequence
64	224	29.4	292	6	AR478733	AR478733 Sequence
65	224	29.4	292	6	AX366610	AX366610 Sequence
66 c	206.5	27.1	603	6	BV163497	BV163497 RPAMMSEQ0
67 c	206.5	27.1	610	6	BV095123	BV095123 RPAMMSEQ0
68	197	25.9	724	4	MEA005356	AJ005356 Macropus
69	191	25.1	857	5	AY563621	AY563621 Xenopus 1
70	187.5	24.6	963	9	AY038181	AY038181 Homo sapi
71	185.5	24.3	1094	9	BC039173	BC039173 Homo sapi
72	183	24.0	999	9	AY358822	AY358822 Homo sapi
73	182	23.9	1043	10	BC019734	BC019734 Mus muscu
74	180	23.6	743	4	AF275314	AF275314 Trichosur
75	180	23.6	999	9	AF488306	AF488306 Homo sapi
76	180	23.6	1486	3	TTTTTS2	X78981 T.trichiura
77	180	23.6	1805	5	AB076019	AB076019 Tribolodo

```
78   172   22.6   897 9   HS447F331   AL591713   Novel hum
79   171   22.4   2812 3  AF036161    AF036161   Trichuris
80   165.5  21.7  374 3   AX716767    AX716767   Sequence
81   165.5  21.7  375 9   CR456977    CR456977   Homo sapi
82   165.5  21.7  390 6   AX716758    AX716758   Sequence
83   165.5  21.7  486 9   AX716765    AX716765   Sequence
84   165.5  21.7  486 9   AY212888    AY212888   Homo sapi
85   165.5  21.7  500 6   AX588473    AX588473   Sequence
86   165.5  21.7  515 6   AX887924    AX887924   Sequence
87   165.5  21.7  515 6   BD027534    BD027534   Sequence
88   165.5  21.7  566 9   BC046106    BC046106   Homo sapi
89   165.5  21.7  567 6   CQ727374    CQ727374   Sequence
90   165.5  21.7  626 6   BC039816    BC039816   Homo sapi
91   165.5  21.7  634 6   BD203758    BD203758   Human nuc
92   165.5  21.7  634 6   BD222208    BD222208   Human nuc
93   165.5  21.7  634 6   AX014329    AX014329   Sequence
94   165.5  21.7  634 6   AX014913    AX014913   Sequence
95   165.5  21.7  643 6   AX588136    AX588136   Sequence
96   165.5  21.7  643 6   AX616357    AX616357   Sequence
97   165.5  21.7  1077 6   AX716759   AX716759   Sequence
98   165.5  21.7  1098 6   AX716761   AX716761   Sequence
99   164.5  21.6  519 6   AX588474    AX588474   Sequence
100  164.5  21.6  577 6   AX588137    AX588137   Sequence
```

ALIGNMENTS

```
RESULT 1
AX302535
LOCUS       AX302535       399 bp    DNA      linear    PAT 30-NOV-2001
DEFINITION  Sequence 53 from Patent WO0175177.
ACCESSION   AX302535
VERSION     AX302535.1  GI:17383077
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Morin,P.J., Sherman-Baust,C.A., Pizer,E.S. and Hough,C.D.
  TITLE     Tumor markers in ovarian cancer
  JOURNAL   Patent: WO 0175177-A 53 11-OCT-2001;
            THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)
FEATURES             Location/Qualifiers
     source          1..399
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                 399
Score:                     762.00
Percent Similarity:        100.00%        Matches:        132
Best Local Similarity:     100.00%        Conservative:   0
Query Match:               100.00%        Mismatches:     0
DB:                        6              Indels:         0
                                          Gaps:           0

US-10-613-105-2 (1-132) x AX302535 (1-399)
```

```
Qy     1 MetLysSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGGCTCTGGGAACTCTGGCACCT 60

Qy    21 TrpAlaValGluGlyLysGlyLysSerPheLysAlaGlyValCysProProLysSer 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TGGGCTGTGGAAGGCTTCGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 120

Qy    41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 180

Qy    61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
```

```
Db   181 AAGAGATGTTGTCCTGACACTTGTGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyValCysLeuMetLeuAsn 100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCAACAAGGAGGAGCCTGGGAGTGCCCAGTGACTTATGGCCAAGTGTTTGATGCTTAAC 300

Qy   101 ProProAsnPheCysGluMetAspGlyAspTrpGlnCysGlyLysLysCysCysMetGly 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGAAGTGTTGCATGGGC 360

Qy   121 MetCysGlyLysSerCysValSerProValLysAla 132
           ||||||||||||||||||||||||||||||||||||
Db   361 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 396
```

```
RESULT 2
AF114471
LOCUS       AF114471       399 bp    mRNA      linear    PRI 30-MAY-2002
DEFINITION  Homo sapiens secretory leukocyte proteinase inhibitor precursor,
            mRNA, complete cds.
ACCESSION   AF114471
VERSION     AF114471.1  GI:4378758
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 399)
  AUTHORS   Si-Tahar,M., Merlin,D., Sitaraman,S. and Madara,J.L.
  TITLE     Constitutive and regulated secretion of secretory leukocyte
            proteinase inhibitor by human intestinal epithelial cells
  JOURNAL   Gastroenterology 118 (6), 1061-1071 (2000)
   MEDLINE  20295039
   PUBMED   10833481
REFERENCE   2  (bases 1 to 399)
  AUTHORS   Si-Tahar,M., Merlin,D., Sitaraman,S. and Madara,J.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-1998) Department of Pathology and Clinical
            Medicine, Emory University, 1639 Pierce Drive, Atlanta, GA 30322,
            USA
FEATURES             Location/Qualifiers
     source          1..399
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /cell_line="Caco2-bbe"
                     /cell_type="epithelium"
                     /tissue_type="intestine"
     CDS             1..399
                     /note="SLPI"
                     /codon_start=1
                     /product="secretory leukocyte proteinase inhibitor
                     precursor"
                     /protein_id="AAD19661.1"
                     /db_xref="GI:4378759"
                     /translation="MKSSGLFPFLVLLALGTLAPWAVEGSGKSFKAGVCPPKSAQCL
                     RYKKPECQSDWQCPGKKRCCPDTCGIKCLDPVDTPNPTRRKPGKCPVTYGQCLMLNPP
                     NFCEMDGDWQCKRDLKCCMCGKSCVSPVKA"
ORIGIN

Alignment Scores:
Pred. No.:                 399
Score:                     3.18e-49       Length:         399
Percent Similarity:        762.00         Matches:        132
Best Local Similarity:     100.00%        Conservative:   0
Query Match:               100.00%        Mismatches:     0
DB:                        9              Indels:         0
                                          Gaps:           0

US-10-613-105-2 (1-132) x AF114471 (1-399)

Qy     1 MetLysSerGlyLeuPheProPheLeuValLeuAlaLeuAlaLeuGlyThrLeuAlaPro 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 3

```
LOCUS       AX778065    551 bp    DNA    linear    PAT 14-JUL-2003
DEFINITION  Sequence 222 from Patent WO03039443.
ACCESSION   AX778065
VERSION     AX778065.1  GI:32695059
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS     Haferlach,T., Schoch,C., Kern,W., Kohlmann,A., Schnittger,S.,
            Dugas,M., Eils,R., Brors,B. and Mergenthaler,S.
TITLE       Novel genetic markers for leukemias
JOURNAL     Patent: WO 03039443-A 222 15-MAY-2003;
            Deutsches Krebsforschungszentrum (DE);
            Ludwig-Maximilian-Universitaet Muenchen (DE) ; Haferlach, Torsten,
            PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)
FEATURES    Location/Qualifiers
source      1..551
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:             4.41e-49
Score:                 762.00                     Length:        551
Percent Similarity:    100.00%                    Matches:       132
Best Local Similarity: 100.00%                    Conservative:  0
Query Match:           100.00%                    Mismatches:    0
DB:                                               Indels:        0
                                                  Gaps:          6

US-10-613-105-2 (1-132) x AX778065 (1-551)

Qy    1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro   20
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   14 ATGAAGTCCAGCGGGCCTCTTCCCCCTCCTGTGCTTGCCCTGGGAACTCTGGCACCT     73
Qy   21 TrpAlaValGluGlyLysGlyLysSerPheLysAlaGlyValCysProProLysLysSer   40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74 TGGGCTGTGGAAGGCCTCGGAAAGTCTTCAAAGCTGGAGTCGTCCCTAAGAAATCT     133
Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys   60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  134 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGCGACTGGCAGTGTCCAGGGAAG   193
```

(continued)

```
Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn   80
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  194 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGATCCTGTTGACACCCCAAAC   253
Qy   81 ProThrArgArgLysProGlyLysProValThrTyrGlyValThrProAsnValMetLeuAsn  100
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  254 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTGATGTTTAAC   313
Qy  101 ProProAsnPheCysGluMetAspGlyGlnLysCysArgAspLeuLysCysMetGly   120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  314 CCCCCCAATTTCTGTGAGATGATGGGCAGTGGCAGGTGCAAGCGTGACTTGAAGTGTGCATGGGC   373
Qy  121 MetCysGlyLysLysSerCysValSerProValLysAla   132
        |||||||||||||||||||||||||||||||||||||
Db  374 ATGTGTGGGAAATCCTGCGTTTCCCGTGTGAAAGCT     409
```

RESULT 4

```
LOCUS       I01502      565 bp    ss-DNA    linear    PAT 21-MAY-1993
DEFINITION  Sequence 3 from Patent US 4845076.
ACCESSION   I01502
VERSION     I01502.1  GI:270158
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 565)
AUTHORS     Heinzel,R., Appelhans,H., Gassen,H.G., Machleidt,W. and
            Seemuller,U.
TITLE       DNA sequences coding for proteins having the biological activity of
            HUSI-type I inhibitors, biotechnological methods for the
            preparation of said proteins and pharmaceutical compositions
            containing said proteins
JOURNAL     Patent: US 4845076-A 3 04-JUL-1989;
            Grunenthal GmbH;;
            DE;
FEATURES    Location/Qualifiers
source      1..565
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:             4.52e-49
Score:                 762.00                     Length:        565
Percent Similarity:    100.00%                    Matches:       132
Best Local Similarity: 100.00%                    Conservative:  0
Query Match:           100.00%                    Mismatches:    0
DB:                                               Indels:        0
                                                  Gaps:          6

US-10-613-105-2 (1-132) x I01502 (1-565)

Qy    1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro   20
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   59 ATGAAGTCCAGCGGGCCTCTTCCCCCTCCTGTGCTTGCCCTGGGAACTCTGGCACCT     118
Qy   21 TrpAlaValGluGlyLysGlyLysSerPheLysAlaGlyValCysProProLysLysSer   40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  119 TGGGCTGTGGAAGGCCTCGGAAAGTCTTCAAAGCTGGAGTCGTCCCTAAGAAATCT     178
Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys   60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  179 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGCGACTGGCAGTGTCCAGGGAAG   238
Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn   80
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  239 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGATCCTGTTGACACCCCAAAC   298
Qy   81 ProThrArgArgLysProGlyLysProValThrTyrGlyValThrProAsnValMetLeuAsn  100
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  299 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTGATGTTTAAC   358
```

```
Qy    101  ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    359  CCCCCAATTTCTGTGAGATGATGGCCAGTCAAGCGTGACTTGAAGTGTTGCATGGGC 418

Qy    121  MetCysGlyLysSerCysValSerProValLysAla 132
           ||||||||||||||||||||||||||||||||||||
Db    419  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 454

RESULT 5
AX069252
LOCUS          AX069252                   565 bp    DNA         linear   PAT 25-JAN-2001
DEFINITION     Sequence 1 from Patent WO010199B.
ACCESSION      AX069252
VERSION        AX069252.1  GI:12579133
KEYWORDS       .
SOURCE         Homo sapiens (human)
  ORGANISM     Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1
  AUTHORS      Neilson,L. and Li,Z.
  TITLE        Treatment of endometriosis with antileukoprotease
  JOURNAL      Patent: WO 010199B-A 1 11-JAN-2001;
               Reprogen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..565
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             59..457
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC27294.1"
                     /db_xref="GI:12579134"
                     /translation="MKSSGLFPPLVLLALGTLAPWAVEGSGKSFKAGVCPPKSAQCL
                     RYKKPECQSDWQCPGKKRCCPDTCGIKCLDPVDTPNPTRRKPGKCPVTYGQCLMLNPP
                     NFCEMDGQCKRDLKCCMGMCGKSCVSPVKA"

ORIGIN

Alignment Scores:
Pred. No.:                4.52e-49       Length:            565
Score:                    762.00         Matches:           132
Percent Similarity:       100.00%        Conservative:      0
Best Local Similarity:    100.00%        Mismatches:        0
Query Match:              100.00%        Indels:            0
DB:                                      Gaps:              6

US-10-613-105-2 (1-132) x AX069252 (1-565)

Qy      1  MetLysSerSerGlyLeuPheProLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     59  ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTGGCCCTAGGAACTCTGGCACCT 118

Qy     21  TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysSer 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    119  TGGGCTGTGGAAGGCTCTGGAAAGTCTTCAAAGCTGGAGTCTGTCTCCTAAGAAATCT 178

Qy     41  AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    179  GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 238

Qy     61  LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    239  AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 298

Qy     81  ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    299  CCAACAAGGAGGAAGCCTGGAAGTGCCCAGTGACCTATGGCCAATGTTTGATGCTTAAC 358

Qy    101  ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    359  CCCCCAATTTCTGTGAGATGATGGCCAGTCAAGCGTGACTTGAAGTGTTGCATGGGC 418
```

```
Qy    121  MetCysGlyLysSerCysValSerProValLysAla 132
           ||||||||||||||||||||||||||||||||||||
Db    419  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 454

RESULT 6
HSSLIPR
LOCUS          HSSLIPR                    573 bp    mRNA        linear   PRI 21-MAR-1995
DEFINITION     Human SLPI mRNA fragment for secretory leucocyte protease
               inhibitor.
ACCESSION      X04503
VERSION        X04503.1  GI:36490
KEYWORDS       elastase inhibitor; protease inhibitor; secretory leucocyte
               protease inhibitor; trypsin inhibitor.
SOURCE         Homo sapiens (human)
  ORGANISM     Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1  (bases 1 to 573)
  AUTHORS      Steler,G., Brewer,M.T. and Thompson,R.C.
  TITLE        Isolation and sequence of a human gene encoding a potent inhibitor
               of leukocyte proteases
  JOURNAL      Nucleic Acids Res. 14 (20), 7883-7896 (1986)
  MEDLINE      87040761
  PUBMED       3640338
COMMENT        Data kindly reviewed (15-SEP-1987) by Steler G.
FEATURES             Location/Qualifiers
     source          1..573
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="parotid gland"
     CDS             1..399
                     /note="unnamed protein product; SLPI-precursor"
                     /codon_start=1
                     /protein_id="CAA28188.1"
                     /db_xref="GI:36491"
                     /db_xref="GOA:P03973"
                     /db_xref="Swiss-Prot:P03973"
                     /translation="MKSSGLFPPLVLLALGTLAPWAVEGSGKSFKAGVCPPKSAQCL
                     RYKKPECQSDWQCPGKKRCCPDTCGIKCLDPVDTPNPTRRKPGKCPVTYGQCLMLNPP
                     NFCEMDGQCKRDLKCCMGMCGKSCVSPVKA"
     sig_peptide     1..75
                     /note="signal peptide (AA -25 to -1)"
     mat_peptide     76..396
                     /product="mature SLPI (AA 1-107)"
     misc_feature    546..552
                     /note="put. polyA signal"
     polyA_site      573
                     /note="polyA site"

ORIGIN

Alignment Scores:
Pred. No.:                4.59e-49       Length:            573
Score:                    762.00         Matches:           132
Percent Similarity:       100.00%        Conservative:      0
Best Local Similarity:    100.00%        Mismatches:        0
Query Match:              100.00%        Indels:            0
DB:                                      Gaps:              9

US-10-613-105-2 (1-132) x HSSLIPR (1-573)

Qy      1  MetLysSerSerGlyLeuPheProLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTGGCCCTGGGAACTCTGGCACCT 60

Qy     21  TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysSer 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCGTCTCCTAAGAAATCT 120

Qy     41  AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 180
```

```
Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
Db  181 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 240

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
Db  241 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 300

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly 120
Db  301 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCCGTGACTTGAGTGTTGCATGGGC 360

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
Db  361 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 396
```

RESULT 7
```
LOCUS       CQ727371                   594 bp   DNA      linear    PAT 03-FEB-2004
DEFINITION  Sequence 13305 from Patent WO02068579.
ACCESSION   CQ727371
VERSION     CQ727371.1   GI:42293168
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE       Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
JOURNAL     Patent: WO 02068579-A 13305 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1..594
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                                 Length:            594
Score:                  4.76e-49           Matches:           132
Percent Similarity:     100.00%            Conservative:        0
Best Local Similarity:  100.00%            Mismatches:          0
Query Match:                               Indels:              0
DB:                     6                   Gaps:               0

US-10-613-105-2 (1-132) x CQ727371 (1-594)

Qy    1 MetLysSerSerGlyLeuPheProPheLeuValLeuValAlaLeuAlaLeuGlyThrLeuAlaPro 20
Db   19 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTCGTGCTGGTGGCCCTGGGAACTCTGGCACCT 78

Qy   21 TrpAlaValGluGluGlyLysGlySerGlyLysSerPheLysAlaGlyValCysProLysSer 40
Db   79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTTAAGAAATCT 138

Qy   41 AlaGlnCysLeuArgTyrArgLysProGluCysGlnSerArgAspTrpGlyProGlyLys 60
Db  139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTGTCCAGAGTGACTGGAG 198

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
Db  199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
Db  259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly 120
```

```
Db  319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCCGTGACTTGAGTGTTGCATGGGC 378

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
Db  379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414
```

RESULT 8
```
LOCUS       AR270708                   594 bp   DNA      linear    PAT 10-APR-2003
DEFINITION  Sequence 1271 from patent US 6500938.
ACCESSION   AR270708
VERSION     AR270708.1   GI:29701942
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 594)
AUTHORS     Au-Young,J. and Seilhamer,J.J.
TITLE       Composition for the detection of signaling pathway gene expression
JOURNAL     Patent: US 6500938-A 1271 31-DEC-2002;
FEATURES             Location/Qualifiers
     source          1..594
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                                 Length:            594
Score:                  4.76e-49           Matches:           132
Percent Similarity:     100.00%            Conservative:        0
Best Local Similarity:  100.00%            Mismatches:          0
Query Match:                               Indels:              0
DB:                     6                   Gaps:               0

US-10-613-105-2 (1-132) x AR270708 (1-594)

Qy    1 MetLysSerSerGlyLeuPheProPheLeuValLeuValAlaLeuAlaLeuGlyThrLeuAlaPro 20
Db   19 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTCGTGCTGGTGGCCCTGGGAACTCTGGCACCT 78

Qy   21 TrpAlaValGluGluGlyLysGlySerGlyLysSerPheLysAlaGlyValCysProLysSer 40
Db   79 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTTAAGAAATCT 138

Qy   41 AlaGlnCysLeuArgTyrArgLysProGluCysGlnSerArgAspTrpGlyProGlyLys 60
Db  139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTGTCCAGAGTGACTGGAG 198

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
Db  199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
Db  259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly 120
Db  319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCCGTGACTTGAGTGTTGCATGGGC 378

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
Db  379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414
```

RESULT 9
```
LOCUS       AX328399                   594 bp   DNA      linear    PAT 07-JAN-2002
DEFINITION  Sequence 1 from Patent WO0190421.
ACCESSION   AX328399
VERSION     AX328399.1   GI:18098360
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
```

REFERENCE   1
AUTHORS   Rastelli,L. and Smithson,G.
TITLE   Method of detecting inflammatory lung disorders
JOURNAL   Patent: WO 0190421-A 1 29-NOV-2001;
            Curagen Corporation (US)
FEATURES   Location/Qualifiers
source   1..594
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.:   594
Score: 4.76e-49   Length: 594
Percent Similarity: 762.00   Matches: 132
Best Local Similarity: 100.00%   Conservative: 0
Query Match: 100.00%   Mismatches: 0
DB: Indels: 0
                                    Gaps: 6

US-10-613-105-2 (1-132) x AX328399 (1-594)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCT 78

Qy   21 TrpAlaValGluGlyLysSerPheLysAlaGlyValCysProProLysLysSer 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGGCTGTGGAAGGCTCTGGAAGGCTCTGGAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 138

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy   61 LysArgCysCysProAspThrCysGlyIleLeuCysProValAspThrProAsn 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysProValThrTyrGlyLysMetLeuAsn 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   259 CCAACAAGGAGGAAGCCTGGGAAGTGCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy   101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGGTGACTTGAAGTGTTGCATGGGC 378

Qy   121 MetCysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||||||||||||
Db   379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414

RESULT 10
AX334507
LOCUS   AX334507   594 bp   DNA   linear   PAT 09-JAN-2002
DEFINITION Sequence 5016 from Patent WO0194629.
ACCESSION   AX334507
VERSION   AX334507.1 GI:18125226
KEYWORDS   .
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE   Cancer gene determination and therapeutic screening using signature
            gene sets
JOURNAL   Patent: WO 0194629-A 5016 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES   Location/Qualifiers
source   1..594
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN

REFERENCE   1
AUTHORS   Rastelli,L. and Smithson,G.
TITLE   Method of detecting inflammatory lung disorders
JOURNAL   Patent: WO 0190421-A 1 29-NOV-2001;
            Curagen Corporation (US)
FEATURES   Location/Qualifiers
source   1..594
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.:   594
Score: 4.76e-49   Length: 594
Percent Similarity: 762.00   Matches: 132
Best Local Similarity: 100.00%   Conservative: 0
Query Match: 100.00%   Mismatches: 0
DB: Indels: 0
                                    Gaps: 6

US-10-613-105-2 (1-132) x AX334507 (1-594)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTTGCCCTGGGAACTCTGGCACCT 78

Qy   21 TrpAlaValGluGlyLysSerPheLysAlaGlyValCysProProLysLysSer 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGGCCTGTGGAAGGCTCTGGAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 138

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 198

Qy   61 LysArgCysCysProAspThrCysGlyIleLeuCysProValAspThrProAsn 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysProValThrTyrGlyLysMetLeuAsn 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   259 CCAACAAGGAGGAAGCCTGGGAAGTGCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 318

Qy   101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGGTGACTTGAAGTGTTGCATGGGC 378

Qy   121 MetCysGlyLysSerCysValSerProValLysAla 132
          ||||||||||||||||||||||||||||||||||||
Db   379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414

RESULT 11
AX335376
LOCUS   AX335376   594 bp   DNA   linear   PAT 09-JAN-2002
DEFINITION Sequence 5885 from Patent WO0194629.
ACCESSION   AX335376
VERSION   AX335376.1 GI:18126095
KEYWORDS   .
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE   Cancer gene determination and therapeutic screening using signature
            gene sets
JOURNAL   Patent: WO 0194629-A 5885 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES   Location/Qualifiers
source   1..594
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.:   594
Score: 4.76e-49   Length: 594
Percent Similarity: 762.00   Matches: 132
Best Local Similarity: 100.00%   Conservative: 0
Query Match: 100.00%   Mismatches: 0
                                    Indels: 0

DB: 6   Gaps: 0

US-10-613-105-2 (1-132) x AX335376 (1-594)

```
Qy    1 MetLysSerGlyLeuPheProPheValLeuLeuAlaLeuGlyThrLeuAlaPro 20
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 ATGAAGTCCAGCGGCCTCTTCCCCTTCGTGCTGCTTGCCCTGGGAACTCTGGCACCT 78

Qy   21 TrpAlaValGluGlyLysSerGlyLysSerPheLysAlaGlyValCysProLysLysSer 40
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGGCTGTGGAAGGCTCTGGAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 138

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerSerAspTrpGlnProGlyLys 60
        |||||||||||||||||||||||||||||||||||||||||||||
Db  139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTCCAGTGACTGGCAGTCCAGGGAAG 198

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
        ||||||||||||||||||||||||||||||||||||||||||||
Db  199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnLeuMetLeuAsn 100
        |||||||||||||||||||||||||||||||||||||||||||
Db  259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTGATGCTTAAC 318

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysMetGly 120
        ||||||||||||||||||||||||||||||||||||||||||
Db  319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTATGGGC 378

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
        ||||||||||||||||||||||||||||||
Db  379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414
```

RESULT 12
AX577960
LOCUS   AX577960   594 bp   DNA   linear   PAT 08-JAN-2003
DEFINITION Sequence 82 from Patent WO02081745.
ACCESSION   AX577960
VERSION   AX577960.1 GI:27647168
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
  AUTHORS   Garcia,T., roman Roman,S., Baron,R., Call,K., Theilhaber,J.,
          Connolly,T., Jackson,A., Bushnell,S.E. and Rawadi,G.
  TITLE   Genes involved in osteogenesis, and methods of use
  JOURNAL   Patent: WO 02081745-A 82 17-OCT-2002;
          Aventis Pharma S.A. (FR)
FEATURES   Location/Qualifiers
    source   1..594
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"
ORIGIN

Alignment Scores:

Pred. No.:	4.76e-49	Length:	594
Score:	762.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:		Indels:	0
DB:	6	Gaps:	0

US-10-613-105-2 (1-132) x AX577960 (1-594)

```
Qy    1 MetLysSerGlyLeuPheProPheValLeuLeuAlaLeuGlyThrLeuAlaPro 20
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 ATGAAGTCCAGCGGCCTCTTCCCCTTCGTGCTGCTTGCCCTGGGAACTCTGGCACCT 78

Qy   21 TrpAlaValGluGlyLysSerGlyLysSerPheLysAlaGlyValCysProLysLysSer 40
        |||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGGCTGTGGAAGGCTCTGGAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 138
```

```
Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerSerAspTrpGlnProGlyLys 60
        ||||||||||||||||||||||||||||||||||||||||||||
Db  139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTCCAGTGACTGGCAGTCCAGGGAAG 198

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
        ||||||||||||||||||||||||||||||||||||||||||||
Db  199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnLeuMetLeuAsn 100
        |||||||||||||||||||||||||||||||||||||||||||
Db  259 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTGATGCTTAAC 318

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysMetGly 120
        ||||||||||||||||||||||||||||||||||||||||||
Db  319 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 378

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
        ||||||||||||||||||||||||||||||
Db  379 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 414
```

RESULT 13
AX780010
LOCUS   AX780010   594 bp   DNA   linear   PAT 14-JUL-2003
DEFINITION Sequence 2167 from Patent WO03039443.
ACCESSION   AX780010
VERSION   AX780010.1 GI:32697004
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
  AUTHORS   Haferlach,T., Schoch,C., Kern,W., Kohlmann,A., Schnittger,S.,
          Dugas,M., Eils,R., Brors,B. and Mergenthaler,S.
  TITLE   Novel genetic markers for leukemias
  JOURNAL   Patent: WO 03039443-A 2167 15-MAY-2003;
          Deutsches Krebsforschungszentrum (DE) ;
          Ludwig-Maximilian-Universitaet Muenchen (DE) ; Haferlach, Torsten,
          PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)
FEATURES   Location/Qualifiers
    source   1..594
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"
ORIGIN

Alignment Scores:

Pred. No.:	4.76e-49	Length:	594
Score:	762.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:		Indels:	0
DB:	6	Gaps:	0

US-10-613-105-2 (1-132) x AX780010 (1-594)

```
Qy    1 MetLysSerGlyLeuPheProPheValLeuLeuAlaLeuGlyThrLeuAlaPro 20
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 ATGAAGTCCAGCGGCCTCTTCCCCTTCGTGCTGCTTGCCCTGGGAACTCTGGCACCT 78

Qy   21 TrpAlaValGluGlyLysSerGlyLysSerPheLysAlaGlyValCysProLysLysSer 40
        |||||||||||||||||||||||||||||||||||||||||||||||
Db   79 TGGGCTGTGGAAGGCTCTGGAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 138

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerSerAspTrpGlnProGlyLys 60
        ||||||||||||||||||||||||||||||||||||||||||||
Db  139 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGTCCAGTGACTGGCAGTCCAGGGAAG 198

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
        |||||||||||||||||||||||||||||||||||||||||||
Db  199 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 258
```

Left column (RESULT 14)

```
Qy   81  ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn  100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259  CCAACAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC  318

Qy  101  ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319  CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC  378

Qy  121  MetCysGlyLysSerCysValSerProValLysAla  132
         ||||||||||||||||||||||||||||||||||||
Db  379  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  414
```

RESULT 14
HSALPR

```
LOCUS       HSALPR                   594 bp    mRNA    linear   PRI 21-MAR-1995
DEFINITION  Human mRNA for antileukoprotease (ALP) from cervix uterus.
ACCESSION   X04470
VERSION     X04470.1  GI:28638
KEYWORDS    antileukoprotease; elastase inhibitor; protease; signal peptide.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 594)
AUTHORS     Heinzel,R., Appelhans,H., Gassen,G., Seemuller,U., Machleidt,W.,
            Fritz,H. and Steffens,G.
TITLE       Molecular cloning and expression of cDNA for human
            antileukoprotease from cervix uterus
JOURNAL     Eur. J. Biochem. 160 (1), 61-67 (1986)
MEDLINE     87030258
PUBMED      3533531
COMMENT     Data kindly reviewed (05-DEC-1986) by H. Appelhans.
FEATURES             Location/Qualifiers
     source          1..594
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             19..417
                     /note="unnamed protein product; precursor ALP"
                     /codon_start=1
                     /protein_id="CAA28158.1"
                     /db_xref="GI:28639"
                     /db_xref="GOA:P03973"
                     /db_xref="Swiss-Prot:P03973"
                     /translation="MKSSGLFPFLVLLALGTLAPWAVEGSGKSFKAGVCPPKKSAQCL
                     RYKKPECQSDWQCPGKKRCCPDTCGIKCLDPVDTPNPTRRKPGKCPVTYGQCLMLNPP
                     NFCEMDGQCKRDLKCCMCGKSCVSPVKA"
     sig_peptide     19..93
                     /note="put. signal peptide (aa -25 to -1)"
     mat_peptide     94..414
                     /product="put. mature peptide (aa 1-107)"
     misc_feature    564..569
                     /note="pot. polyA signal"
     misc_feature    568..573
                     /note="pot. polyA signal"
ORIGIN

Alignment Scores:
Pred. No.:
Score:                  4.76e-49        Length:          594
                        762.00          Matches:         132
Percent Similarity:     100.00%         Conservative:    0
Best Local Similarity:  100.00%         Mismatches:      0
Query Match:                            Indels:          0
DB:                     9               Gaps:            0

US-10-613-105-2 (1-132) x HSALPR (1-594)

Qy    1  MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19  ATGAAGAGTTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTCGCCCTTGGAACTCTGGCACCT  78

Qy   21  TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer  40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Right column (RESULT 15)

```
Db   79  TGGGCTGTGGAAGGCTCTCGGAAGTCCTTCAAAGCTGGAGTGTCTCCTCCTTCAAAGAAATCT  138

Qy   41  AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  139  GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG  198

Qy   61  LysArgCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  199  AAGAGATGTGTCTGCACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC  258

Qy   81  ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn  100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259  CCAACAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC  318

Qy  101  ProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLeuLysCysCysMetGly  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319  CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC  378

Qy  121  MetCysGlyLysSerCysValSerProValLysAla  132
         ||||||||||||||||||||||||||||||||||||
Db  379  ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT  414
```

RESULT 15
BD203743

```
LOCUS       BD203743                 599 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Human nucleic acid sequence originating in ovarian myoma tissue.
ACCESSION   BD203743
VERSION     BD203743.1  GI:33013513
KEYWORDS    JP 200251252-A/84.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 599)
AUTHORS     Specht,T., Hinzmann,B., Schmitt,A., Pilarsky,C., Dahl,E. and
            Rosenthal,A.
TITLE       Human nucleic acid sequence originating in ovarian myoma tissue
JOURNAL     Patent: JP 200251252-A 84 16-APR-2002;
            METAGEN GESELLSCHAFT FUER GENOME FORSCHUNG MBH
COMMENT     OS   Homo sapiens (human)
            PN   JP 200251252-A/84
            PD   16-APR-2002
            PF   07-APR-1999 JP 2000543588
            PR   09-APR-1998 DE 198 17 557.4
            PI   THOMAS SPECHT,BERND HINZMANN,ARMIN SCHMITT,CHRISTIAN PILARSKY,
            PI   EDGAR DAHL,
            PI   ANDRE ROSENTHAL
            PC   C12N15/09,A61K38/00,A61K48/00,A61P15/00,A61P35/00,C07K14/82,
            PC   C07K16/32,
            PC   C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/
            PC   574,
            PC   C12N15/00,A61K37/02,C12N5/00
            CC   Human nucleic acid sequence originating in ovarian myoma
            CC   tissue
FEATURES             Location/Qualifiers
     source          1..599
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:
Score:                  4.8e-49         Length:          599
                        762.00          Matches:         132
Percent Similarity:     100.00%         Conservative:    0
Best Local Similarity:  100.00%         Mismatches:      0
Query Match:                            Indels:          0
DB:                     6               Gaps:            0

US-10-613-105-2 (1-132) x BD203743 (1-599)
```

```
Qy      1 MetLysSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     14 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTGGCCCTGGGAACTCTGGCACCT 73

Qy     21 TrpAlaValGluLysSerGlyLySerPheLysAlaGlyValCysProProLysLysSer 40
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74 TGGGCTGTGGAGAAGTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 133

Qy     41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    134 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 193

Qy     61 LysArgCysAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn 80
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    194 AAGAGATGTGTCCTGACACTTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 253

Qy     81 ProThrArgArgLysProGlyLysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    254 CCAACAAGGAGGAAGCCTGGAAAGCCTGTGACTTATGGCCAATGTTTGATGCTTAAC 313

Qy    101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    314 CCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGC 373

Qy    121 MetCysGlyLysSerCysValSerProValLysAla 132
            ||||||||||||||||||||||||||||||||||||
Db    374 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT 409
```

RESULT 17
BC020708
LOCUS   BC020708   625 bp   mRNA   linear   PRI 29-JUN-2004
DEFINITION   Homo sapiens secretory leukocyte protease inhibitor
             (antileukoproteinase), mRNA (cDNA clone MGC:22479 IMAGE:4733996),
             complete cds.
ACCESSION   BC020708
VERSION   BC020708.1   GI:18008404
KEYWORDS   MGC.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 625)
AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
             Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
             Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
             Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
             Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
             Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
             Scheetz,T.B., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
             Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
             Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
             McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
             Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
             Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
             Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
             Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
             Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
             Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
             Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
             Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE   Generation and initial analysis of more than 15,000 full-length
             human and mouse cDNA sequences
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED   12477932
REFERENCE   2 (bases 1 to 625)
AUTHORS   Strausberg,R.
TITLE   Direct Submission
JOURNAL   Submitted (03-JAN-2002) National Institutes of Health, Mammalian
             Gene Collection (MGC), Cancer Genomics Office, National Cancer
             Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
             USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: CLONTECH
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Sequencing Group at the Stanford Human Genome
             Center, Stanford University School of Medicine, Stanford, CA 94305
             Web site: http://www-shgc.stanford.edu
             Contact: (Dickson, Mark) mcd@paxil.stanford.edu

RESULT 16
AX014898
LOCUS   AX014898   599 bp   DNA   linear   PAT 07-SEP-2000
DEFINITION   Sequence 97 from Patent WO9953040.
ACCESSION   AX014898
VERSION   AX014898.1   GI:10041165
KEYWORDS   .
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS   Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
             Pilarsky,C.
TITLE   Human nucleic acid sequences from ovarian tumour tissue
JOURNAL   Patent: WO 9953040-A 97 21-OCT-1999;
             SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
             BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
             (DE); PILARSKY CHRISTIAN (DE)
FEATURES   Location/Qualifiers
     source 1..599
             /organism="Homo sapiens"
             /mol_type="unassigned DNA"
             /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:   4.8e-49
Score: 762.00
Percent Similarity: 100.00%   Length: 599
Best Local Similarity: 100.00%   Matches: 132
Query Match: 100.00%   Conservative: 0
DB: Mismatches: 0
                                         Indels: 0
                                         Gaps: 6
```

US-10-613-105-2 (1-132) x AX014898 (1-599)

```
Qy      1 MetLysSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     14 ATGAAGTCCAGCGGCCTCTTCCCCTTCCTGGTGCTGCTGGCCCTGGGAACTCTGGCACCT 73

Qy     21 TrpAlaValGluLysSerGlyLySerPheLysAlaGlyValCysProProLysLysSer 40
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74 TGGGCTGTGGAGAAGTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 133
```

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 37 Row: i Column: 13
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 15834622.

```
FEATURES             Location/Qualifiers
     source          1..625
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:22479 IMAGE:4733996"
                     /tissue_type="Liver"
                     /clone_lib="NIH_MGC_76"
                     /lab_host="DH10B"
                     /note="Vector: pDNR-LIB"
     gene            1..625
                     /gene="SLPI"
                     /note="synonyms: HUSI, WAP4, WFDC4, ALK1, ALP, BLPI,
                     HUSI-I"
                     /db_xref="LocusID:6590"
                     /db_xref="MIM:107285"
     CDS             22..420
                     /gene="SLPI"
                     /codon_start=1
                     /product="secretory leukocyte protease inhibitor,
                     precursor"
                     /protein_id="AAH20708.1"
                     /db_xref="GI:18088405"
                     /db_xref="LocusID:6590"
                     /db_xref="MIM:107285"
                     /translation="MKSSGLFPFIVLLALGTLAPWAVEGSGKSFKAGVCPPKKSAQCL
                     RYKKPECQSDWQCPGKKRCCPDTCGIKCLDPVDTPNPTRRKPGKCPVTYGQCLMLNPP
                     NFCEMDGQCKRDLKCCMGMCGKSCVSPVKA"
```

ORIGIN
Alignment Scores:
Pred. No.:   625
Score: 5.01e-49
Percent Similarity: 762.00
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 9

Length: 625
Matches: 132
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-613-105-2 (1-132) x BC020708 (1-625)

```
Qy    1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
Db   22 ATGAAGTCCAGCGGCCGCCTCTTCCCGTTCCTGGTGCTTCTGGCGCTTGGAACTCTGGCACCT  81

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer  40
Db   82 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT  141

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
Db  142 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG  201

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
Db  202 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTAGATCCTGTTGACACCCCAAAC  261

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
Db  262 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC  321

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
Db  322 CCCCCAATTTCTGTGAGATGGATGGCCAAGCGTGACTTGAAGTGTTGCATGGGC  381

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
```

```
RESULT 18
CQ413516
LOCUS       CQ413516          1084 bp   DNA   linear   PAT 23-JAN-2004
DEFINITION  Sequence 20587 from Patent WO0170979.
ACCESSION   CQ413516
VERSION     CQ413516.1  GI:41321297
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lee,J. and Lillie,J.
  TITLE     Genes, compositions, kits, and method for identification,
            assessment, prevention, and therapy of ovarian cancer
  JOURNAL   Patent: WO 0170979-A 20587 27-SEP-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..1084
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

Alignment Scores:
Pred. No.:   1084
Score: 8.74e-49
Percent Similarity: 762.00
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 6

Length: 1084
Matches: 132
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-613-105-2 (1-132) x CQ413516 (1-1084)

```
Qy    1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro  20
Db   69 ATGAAGTCCAGCGGCCGCCTCTTCCCGTTCCTGGTGCTTCTGGCGCTTGGAACTCTGGCACCT 128

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer  40
Db  129 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCTGGAGTCTGTCCTCCTAAGAAATCT 188

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLys  60
Db  189 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG 248

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn  80
Db  249 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC 308

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn 100
Db  309 CCAACAAGGAGGAAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATGCTTAAC 368

Qy  101 ProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGly 120
Db  369 CCCCCAATTTCTGTGAGATGGATGGCCAAGCGTGACTTGAAGTGTTGCATGGGC 428

Qy  121 MetCysGlyLysSerCysValSerProValLysAla 132
Db  429 ATGTGTGGAAATCCTGCGTTTCCCTGTGAAAGCT 464
```

```
RESULT 19
AX772818
LOCUS       AX772818          605 bp   mRNA   linear   PAT 09-JUL-2003
DEFINITION  Sequence 15 from Patent WO03046180.
ACCESSION   AX772818
VERSION     AX772818.1  GI:32485045
KEYWORDS    .
SOURCE      Homo sapiens (human)
```

ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
  AUTHORS Bejanin,S. and Tanaka,H.
  TITLE   Human cDNAs and proteins and uses thereof
  JOURNAL Patent: WO 0304180-A 15 05-JUN-2003;
          Genset S.A. (FR)
FEATURES Location/Qualifiers
  source 1..605
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
  5'UTR   1..21
  CDS   22..363
          /note="unnamed protein product"
          /codon_start=1
          /protein_id="CAE06258.1"
          /db_xref="GI:32485046"
          /translation="MKSSGLFPFIVLLALGTLAPWAVEGSGKSFKAGVCPPKKSAQCL
          RYKKPECQSDWQCEGKKRCCPDTCGIKCLDPVDTPNPTRRKPGKCPVTYGQCLMLNPP
          NPCEMDGQCVT"
  3'UTR   361..605
  polyA_signal 565..570
  polyA_site   590..605
ORIGIN

Alignment Scores:
Pred. No.:   6.62e-46   Length: 605
Score: 720.50   Matches: 129
Percent Similarity: 97.73%   Conservative: 0
Best Local Similarity: 97.73%   Mismatches: 2
Query Match: 94.55%   Indels: 2
DB: Gaps: 1

US-10-613-105-2 (1-132) x AX772818 (1-605)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro   20
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22 ATGAAGTCCAGCGGCCTCTTCCCCTTCCCCTTCGTGCTGCTCGCCCTGGGAACTCTGGCACCT   81

Qy   21 TrpAlaValGluGlySerGlyLysSerPheLysAlaGlyValCysProProLysLysSer   40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82 TGGGCTGTGGAAGGCTCTGGAAAGTCCTTCAAAGCCGGAGTCTGTCCTCCTAAGAAATCT   141

Qy   41 AlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysGlyProGlyLys   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142 GCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGTCCAGGGAAG   201

Qy   61 LysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsn   80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 202 AAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAAC   261

Qy   81 ProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsn   100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 262 CCAACAAGGAGGAAGCCCGGAAGTCCCAGTGACTTATGCCAATGTTTGATGCTTAAC   321

Qy 101 ProProAsnPheCysGluMetAspGlyGlnCysValSerArgLysCysCysMetGly   120
         |||||||||||||||||||||||||||||||||||| ||||||||||
Db 322 CCCCCAAATTCTGTGAGATGGATGGCCAGTGCGT-TAC---TTGAAGTGTTGCATGGGC   377

Qy 121 MetCysGlyLysSerCysValSerProValLysAla   132
         ||||||||||||||||||||||||||||||||||
Db 378 ATGTGTGGGAAATCCTGCGTTTCCCCTGTGAAAGCT   413

RESULT 20
AX670655
LOCUS   AX670655   1525 bp   DNA   linear   PAT 26-MAR-2003
DEFINITION Sequence 15 from Patent WO0250287.
ACCESSION   AX670655
VERSION   AX670655.1 GI:29292050
KEYWORDS
SOURCE   Homo sapiens (human)

ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
  AUTHORS Barr,P.J., Gibson,H.L. and Pemberton,P.
  TITLE   Multifunctional protease inhibitors and their use in treatment of
          disease
  JOURNAL Patent: WO 0250287-A 15 27-JUN-2002;
          Arriva Pharmaceuticals, Inc. (US)
FEATURES Location/Qualifiers
  source 1..1525
          /organism="Homo sapiens"
          /mol_type="unassigned DNA"
          /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:   2.64e-39   Length: 1525
Score: 638.50   Matches: 112
Percent Similarity: 87.97%   Conservative: 5
Best Local Similarity: 84.21%   Mismatches: 9
Query Match: 83.79%   Indels: 7
DB: Gaps: 1

US-10-613-105-2 (1-132) x AX670655 (1-1525)

Qy   7 PheProPheLeuValLeu------------------LeuAlaLeuGlyThrLeuAla   19
         ||||||||||||::::: ||| :::||||
Db 1119 TTCGTTTTCTGATGATCGAGCAGAACACTAAAGCCCCATTGTTATGGGTAAGGTTGTC   1178

Qy   20 ProTrpAlaValGluLysProGlyLysSerPheAlaGlyValCysProProLysLys   39
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1179 AACCCACTCAGAAGATCCGGAAAGTCTTTCAAGGCCGGTGTTTTGTCCACCAAAGAAG   1238

Qy   40 SerAlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnSerProGly   59
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1239 TCCGCTCAATGTTTGAGATACAAGAAGCCAGAATGTCAATCCGACTGGCAATGTCCAGGT   1298

Qy   60 LysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspThrPro   79
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1299 AAGAAGAGATGTTGTCCAGACACTTGGTATCAAGTGTCTAGACCCCAGTTGACACCCCA   1358

Qy   80 AsnProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeu   99
         ||||||||||||||||||||||| :::|||||||||||||||||||||||||||||
Db 1359 AACCCAACTAGAAGAAAGCCAGGTAAGTGTCCAGTTACTTACGGTCAATGTTTGATGTTG   1418

Qy 100 AsnProProAsnPheCysGluMetAspGlyGlnCysValGlnLysArgArgAspAspLeuLysCysCysMet   119
         |||||||||||||||||||||||||||||||||
Db 1419 AACCCACCAAACTTCTGTGAAATGGACGGTCAATGTGAAAGACTTGAAGTGTTGTATG   1478

Qy 120 GlyMetGlyGlyLysProValSerProValLysAla   132
         |||||||||||||||||||||||||||||||||
Db 1479 GGTATGTGTGGTAAGTCCTGTGTTTCCCCAGTCAAGGCC   1517

RESULT 21
AX670643
LOCUS   AX670643   321 bp   DNA   linear   PAT 26-MAR-2003
DEFINITION Sequence 3 from Patent WO0250287.
ACCESSION   AX670643
VERSION   AX670643.1 GI:29292045
KEYWORDS
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
  AUTHORS Barr,P.J., Gibson,H.L. and Pemberton,P.
  TITLE   Multifunctional protease inhibitors and their use in treatment of
          disease
  JOURNAL Patent: WO 0250287-A 3 27-JUN-2002;
          Arriva Pharmaceuticals, Inc. (US)
FEATURES Location/Qualifiers
  source 1..321

```
                    /organism="Homo sapiens"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                      1e-39            Length:              321
Score:                          635.00           Matches:             107
Percent Similarity:             100.00%          Conservative:        0
Best Local Similarity:          100.00%          Mismatches:          0
Query Match:                    83.33%           Indels:              0
DB:                             6                Gaps:                0

US-10-613-105-2 (1-132) x AX670643 (1-321)

Qy    26  SerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeuArg  45
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  TCTGGAAAGTCTTTCAAGGCCGGTGTTTGTCCACCAAAGAGAGTCCGCTCAATGTTTGAGA  60

Qy    46  TyrLysLysProGluGlnSerAspTrpGlnCysProGlyLysLysLysArgCysCysPro  65
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  TACAAGAAGCCAGAATGTCAATCCGACTGGCAATGTCCAGGTAAGAAGAAGATGTTGTCCA  120

Qy    66  AspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArgArgLys  85
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GACACCTGTGGTATCAAGTGTCTAGACCCAGTTGACACCCCAACTCCAACTAGAAGAAAG  180

Qy    86  ProGlyLysCysProValThrTyrGlyTyrGlyLeuMetLeuAsnProProAsnPheCys  105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  CCAGGTAAGTGCCCAGTTACTTACGGTTATGGTTTGATGTTGAACCCACCAAACTTCTGT  240

Qy   106  GluMetAspGlyGlnCysArgAspLeuLysCysCysGlyMetCysGlyLysLysSer  125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  GAAATGGACGGCCAGTGTAGAGATCTGAAGTGTTGTGGTATGTGTGGTAAGTCC  300

Qy   126  CysValSerProValLysAla  132
          |||||||||||||||
Db   301  TGTGTTTCCCCAGTCAAGGCC  321

RESULT 22
I08404
LOCUS       I08404                    324 bp    DNA    linear    PAT 02-DEC-1994
DEFINITION  Sequence 1 from Patent WO 8603519.
ACCESSION   I08404
VERSION     I08404.1  GI:588883
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 324)
  AUTHORS   Bandyopadhyay,P.K., Eisenberg,S.P., Stetler,G.L. and Thompson,R.C.
  TITLE     RECOMBINANT METHODS FOR PRODUCTION OF SERINE PROTEASE INHIBITORS
            AND DNA SEQUENCES USEFUL FOR SAME
  JOURNAL   Patent: WO 8603519-A 1 19-JUN-1986;
FEATURES             Location/Qualifiers
     source          1..324
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                      1.01e-39         Length:              324
Score:                          635.00           Matches:             107
Percent Similarity:             100.00%          Conservative:        0
Best Local Similarity:          100.00%          Mismatches:          0
Query Match:                    83.33%           Indels:              0
DB:                             6                Gaps:                0

US-10-613-105-2 (1-132) x I08404 (1-324)

Qy    26  SerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeuArg  45
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  AGCGGTAAAAGCTTCAAAGCTGGGTGTATGCCCCGCCCGAAAAAATCCCGCCAGTGTCTGCGG  60
```

```
Qy    46  TyrLysProGlnSerAspTrpGlnCysProGlyLysLysLysArgCysCysPro  65
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  TACAAAAAACCGGAATGCCAGTCCGACTGGCAGTGCCCGGGTAAAAAACGTTGTTGCCCG  120

Qy    66  AspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArgArgLys  85
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GACACCTGCGGCATCAAATGCCTGGATCCGGTTGATACCCCGAACCCGACTCGTCGAAAA  180

Qy    86  ProGlyLysCysProValThrTyrGlyTyrGlyLeuMetLeuAsnProProAsnPheCys  105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  CCGGGTAAATGCCCGGTAACCTACGGTTACGGTCTGATGCTGAACCCGCCGAACTTCTGC  240

Qy   106  GluMetAspGlyGlnCysArgAspLeuLysCysCysGlyMetCysGlyLysLysSer  125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  GAAATGGACGGCCAGTGTAAACGATCTGAAATGTTGTGGCATGTGCGGCAAATCT  300

Qy   126  CysValSerProValLysAla  132
          |||||||||||||||
Db   301  TGTGTTTCCCCGGTAAAAGCA  321

RESULT 23
AX670647
LOCUS       AX670647                 1525 bp    DNA    linear    PAT 26-MAR-2003
DEFINITION  Sequence 7 from Patent WO0250287.
ACCESSION   AX670647
VERSION     AX670647.1  GI:29292047
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Barr,P.J., Gibson,H.L. and Pemberton,P.
  TITLE     Multifunctional protease inhibitors and their use in treatment of
            disease
  JOURNAL   Patent: WO 0250287-A 7 27-JUN-2002;
            Arriva Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..1525
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                      4.85e-39         Length:              1525
Score:                          635.00           Matches:             107
Percent Similarity:             100.00%          Conservative:        0
Best Local Similarity:          100.00%          Mismatches:          0
Query Match:                    83.33%           Indels:              0
DB:                             6                Gaps:                0

US-10-613-105-2 (1-132) x AX670647 (1-1525)

Qy    26  SerGlySerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeuArg  45
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    12  TCTGGAAAGTCTTTCAAGGCCGGTGTTTGTCCACCAAAGAGAGTCCGCTCAATGTTTGAGA  71

Qy    46  TyrLysProGluGlnSerAspTrpGlnCysProGlyLysLysLysArgCysCysPro  65
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    72  TACAAGAAGCCAGAATGTCAATCCGACTGGCAATGTCCAGGTAAGAAGAGATGTTGTCCA  131

Qy    66  AspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArgArgLys  85
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   132  GACACTTGTGGTATCAAGTGTCTAGACCCAGTTGACACCCCAACTCCAACTAGAAGAAAG  191

Qy    86  ProGlyLysCysProValThrTyrGlyTyrGlyLeuMetLeuAsnProProAsnPheCys  105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   192  CCAGGTAAGTGCCCAGTTACTTACGGTTACGGTCTGATGTTGAACCCACCAAACTTCTGT  251

Qy   106  GluMetAspGlyGlnCysArgAspLeuLysCysCysGlyMetCysGlyLysLysSer  125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  
```

Db   252 GAAATGGACGGTCAATGTAAGAGAGACTTGAAGTGTTGTATGGGTATGTGTGGTAAGTCC 311

Qy   126 CysValSerProValLysAla 132
                 |||||||||||||||||||||
Db   312 TGTGTTTCCCAGTCAAGGCC 332

RESULT 24
A08111
LOCUS   A08111   321 bp   DNA   linear   PAT 28-JUL-1993
DEFINITION   Synthetic ALP-gene 242.
ACCESSION   A08111
VERSION   A08111.1   GI:413358
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
                artificial sequences.
REFERENCE   1 (bases 1 to 321)
  AUTHORS   Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
  TITLE   Serine protease inhibitor proteins, medicaments containing them,
                DNA sequences coding for these proteins and methods for producing
                these proteins, medicaments and DNA sequences
  JOURNAL   Patent: EP 0373335-A 33 20-JUN-1990;
                Gruenenthal GmbH
FEATURES   Location/Qualifiers
     source   1..321
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
     gene   1..321
                     /gene="ALP-242"
     CDS   1..321
                     /gene="ALP-242"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA00747.1"
                     /db_xref="GI:413359"
                     /translation="SGKSFKAGVCPPKKSAQCLRYKKPECQSDWQCPGKKRCCPDTCG
                     IKCLDPVDTPNPTRRKPGKCPVTYGQCLLLNPPNFCEMDGQCKRDLKCCMGMCGKSCV
                     SPVKA"
ORIGIN
Alignment Scores:
Pred. No.:   1.69e-39   Length: 321
Score: 632.00   Matches: 106
Percent Similarity: 100.00%   Conservative: 1
Best Local Similarity: 99.07%   Mismatches: 0
Query Match: 82.94%   Indels: 0
DB: 6   Gaps: 0

US-10-613-105-2 (1-132) x A08111 (1-321)

Qy   26 SerGlyLysSerPheLysAlaGlyValCysProProLysLysSerAlaGlnCysLeuArg 45
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 AGCGGTAAGTCTTTCAAAGCCGGCGGGTGTCCTCCTCCTAAGAAATCTGCTCAGTGCCGT 60

Qy   46 TyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysPro 65
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TACAAGAAACCCGAGTGCCAGTCTGACTGGCAGTGTCCTGGTAAGAAGCGTTGTTGTCCT 120

Qy   66 AspThrCysGlyIleLysCysGlnSerAspProValAspThrProAsnProThrArgArgLys 85
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GACACTTGTGGTATCAAATGCCTGGATCCGGTTGACACCCCGAACCCGACGCGTCGTAAA 180

Qy   86 AspThrCysGlyIleLeuLysAspProValAspThrProAsnProThrArgArgLys 85
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GACACTTGTGGTATCAAATGCCTGGATCCGGTTGACACCCCGAACCCGACGCGTCGTAAA 180

Qy   106 GluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMetCysGlyLysSer 125
                 |||||||||||||||| :||||||||||||||||||||||||||||||||||||||
Db   241 GAGATGGATGGCCAGTGCAAACGAGATCTGAAATGCTGCATGGGTATGTGCGGTAAAAGC 300

Qy   126 CysValSerProValLysAla 132
                 |||||||||||

Left column

```
LOCUS       A08107                   321 bp    DNA        linear   PAT 28-JUL-1993
DEFINITION  Synthetic ALP-gene 231.
ACCESSION   A08107
VERSION     A08107.1  GI:413350
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 321)
  AUTHORS   Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
  TITLE     Serine protease inhibitor proteins, medicaments containing them,
            DNA sequences coding for these proteins and methods for producing
            these proteins, medicaments and DNA sequences
  JOURNAL   Patent: EP 0373335-A 29 20-JUN-1990;
            Gruenenthal GmbH
FEATURES             Location/Qualifiers
     source          1..321
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
     gene            1..321
                     /gene="ALP-231"
     CDS             1..321
                     /gene="ALP-231"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA00743.1"
                     /db_xref="GI:413351"
                     /translation="SGKSFKAGVCPPKSAQCLRYKKPECQSDWQCPGKKRCCPDTCG
                     IKCLDPVDTPNPTRRKPGKCPVTYGQCLLLNPPNFCELDGQCKRDLKCCLGLCGKSCV
                     SPVKA"
ORIGIN

Alignment Scores:
Pred. No.:              8.1e-39          Length:           321
Score:                  623.00           Matches:          103
Percent Similarity:     100.00%          Conservative:       4
Best Local Similarity:  96.26%           Mismatches:         0
Query Match:            81.76%           Indels:             0
DB:                                      Gaps:               0

US-10-613-105-2 (1-132) x A08107 (1-321)

Qy    26 SerGlyLysPheLysAlaGlyValCysProProLysSerAlaGlnCysLeuArg 45
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 AGCGGTAAGTCTTTCAAAGCCGGCGGGTCTGTCCTCCTAAGAAATCTGCTCAGTGCCTGCGT 60

Qy    46 TyrLysLysProGluGlnCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysPro 65
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TACAAGAAACCCGAGTGCCAGTCTGACTGGCAGTGTCCTGGTAAGAAGAGGTGTTGTCCT 120

Qy    66 AspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArgArgLys 85
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GACACTTGTGGTATCAAATGCCTGGATCCGGTTGACACCCCGAACCCGACGCGGTCGTAAA 180

Qy    86 ProGlyLysCysProValThrTyrGlyGlnCysLeuLeuLeuAsnProProAsnPheCys 105
         |||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db   181 CCCGGGAAGTGCCCGGTTACCTACGGTCAGTGCCTGCTGCTGAACCCGCCTAACTTCTGC 240

Qy   106 GluLeuAspGlyGlnCysLysArgAspLeuLysCysCysLeuGlyMetCysGlyLysSer 125
         |||||||||||||||||||||||||||||:::||:|||||||||||||||||||||||||
Db   241 GAGCTCGATGGCCAGTGCAAACGAGATCTGAAATGCTGCCTGGGTCTGTGCGGTAAAAGC 300

Qy   126 CysValSerProValLysAla 132
         |||||||||||||||||
Db   301 TGCGTTAGCCCGGTTAAGGCT 321

RESULT 27
A08112
LOCUS       A08112                   321 bp    DNA        linear   PAT 28-JUL-1993
DEFINITION  Synthetic ALP-gene 246.
ACCESSION   A08112
VERSION     A08112.1  GI:413360
```

Right column

```
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 321)
  AUTHORS   Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
  TITLE     Serine protease inhibitor proteins, medicaments containing them,
            DNA sequences coding for these proteins and methods for producing
            these proteins, medicaments and DNA sequences
  JOURNAL   Patent: EP 073335-A 34 20-JUN-1990;
            Gruenenthal GmbH
FEATURES             Location/Qualifiers
     source          1..321
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
     gene            1..321
                     /gene="ALP-246"
     CDS             1..321
                     /gene="ALP-246"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA00748.1"
                     /db_xref="GI:413361"
                     /translation="SGKSFKAGVCPPKKSAQCLRYKKPECQSDWQCPGKKRCCPDTCG
                     IKCLDPVDTPNPTRRKPGKCPVTEGQCLLLNPPNFCEMDGQCKRDLKCCMGMCGKSCV
                     SPVKA"
ORIGIN

Alignment Scores:
Pred. No.:              8.1e-39          Length:           321
Score:                  623.00           Matches:          105
Percent Similarity:     99.07%           Conservative:       1
Best Local Similarity:  98.13%           Mismatches:         1
Query Match:            81.76%           Indels:             0
DB:                                      Gaps:               0

US-10-613-105-2 (1-132) x A08112 (1-321)

Qy    26 SerGlyLysPheLysAlaGlyValCysProProLysSerAlaGlnCysLeuArg 45
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 AGCGGTAAGTCTTTCAAAGCCGGCGGGTCTGTCCTCCTAAGAAATCTGCTCAGTGCCTGCGT 60

Qy    46 TyrLysLysProGluSerAspTrpGlnCysProGlyLysLysArgCysCysPro 65
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TACAAGAAACCCGAGTGCCAGTCTGACTGGCAGTGTCCTGGTAAGAAGAGGTGTTGTCCT 120

Qy    66 AspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArgArgLys 85
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GACACTTGTGGTATCAAATGCCTGGATCCGGTTGACACCCCGAACCCGACGCGGTCGTAAA 180

Qy    86 ProGlyLysCysProValThrTyrGlyGlnCysLeuLeuMetLeuAsnProProAsnPheCys 105
         |||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db   181 CCCGGGAAGTGCCCGGTTACCTACGGTCAGTGCCTGCTGCTGAACCCGCCTAACTTCTGC 240

Qy   106 GluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMetCysGlyLysSer 125
         |||||||||||||||||||||||||||:::||:|||||||||||||||||||||||||||
Db   241 GAGATGGATGGCCAGTGCAAACGAGATCTGAAATGCTGCCTGGGTATGTGCGGTAAAAGC 300

Qy   126 CysValSerProValLysAla 132
         |||||||||||||||||
Db   301 TGCGTTAGCCCGGTTAAGGCT 321

RESULT 28
A08108
LOCUS       A08108                   321 bp    DNA        linear   PAT 28-JUL-1993
DEFINITION  Synthetic ALP-gene 232.
ACCESSION   A08108
VERSION     A08108.1  GI:413352
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
```

REFERENCE   1 (bases 1 to 321)
AUTHORS   Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
TITLE   Serine protease inhibitor proteins, medicaments containing them,
            DNA sequences coding for these proteins and methods for producing
            these proteins, medicaments and DNA sequences
JOURNAL   Patent: EP 0373335-A 30 20-JUN-1990;
            Gruenenthal GmbH
FEATURES   Location/Qualifiers
     source   1..321
                    /organism="synthetic construct"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:32630"
     gene   1..321
                    /gene="ALP-232"
     CDS   1..321
                    /gene="ALP-232"
                    /codon_start=1
                    /transl_table=11
                    /protein_id="CAA00744.1"
                    /db_xref="GI:413355"
                    /translation="SGKSFKAGVCPPKKSAQCLRYKKPEQSDWQCPGKKRCCPDTCG
                    IKCLDPVDTPNPTRRKPGKCPVTYGQCILLNPPNFCELDGQCKRDLKCCLGFCGKSCV
                    SPVKA"
ORIGIN

Alignment Scores:
Pred. No.:   1.15e-38   Length: 321
Score: 621.00   Matches: 102
Percent Similarity: 100.00%   Conservative: 5
Best Local Similarity: 95.33%   Mismatches: 0
Query Match: 81.50%   Indels: 0
DB: 6   Gaps: 0

US-10-613-105-2 (1-132) x A08108 (1-321)

Qy 26 SerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeuArg 45
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 AGCGGTAAGTCTTTCAAAGCCGGCGTCTGTCCTCCTAAGAAATCTGCTCAGTGCCTGCGT 60

Qy 46 TyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysPro 65
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 TACAAGAAACCCGAGTGCCAGTCTGACTGGCAGTGTCCTGGTAAGAAGCGGTGTTGTCCT 120

Qy 66 AspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArgArgLys 85
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GACACTTGTGGTATCAAATGCCTGGATCCGGTTGACACCCCGAACCCGACGCGTCGTAAA 180

Qy 86 ProGlyLysCysProValThrTyrGlyGlnCysLeuLeuLeuAsnProProAsnPheCys 105
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 CCCGGAAGTGCCCGGTTACCTACGGTCAGTGCCTGCTGCTGAACCCGCCTAACTTCTGC 240

Qy 106 GluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyLysLysGlyLysSer 125
          ||||||||||||||||||||||||||| :::::|||||||||||||||||||||||||
Db 241 GAGCTCGATGGCCAGTGCAAACGAGATCTGAAATGCTGCCTGGGTCGTGTGCGGTAAAAGC 300

Qy 126 CysValSerProValLysAla 132
          |||||||||||||||||||||
Db 301 TGCGTTAGCCCGGTTAAGGCT 321

RESULT 29
LOCUS   A08110   321 bp   DNA   linear   PAT 28-JUL-1993
DEFINITION Synthetic ALP-gene 237.
ACCESSION   A08110
VERSION   A08110.1 GI:413356
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 321)
AUTHORS   Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
TITLE   Serine protease inhibitor proteins, medicaments containing them,
            DNA sequences coding for these proteins and methods for producing

---

JOURNAL   these proteins, medicaments and DNA sequences
            Patent: EP 0373335-A 32 20-JUN-1990;
            Gruenenthal GmbH
FEATURES   Location/Qualifiers
     source   1..321
                    /organism="synthetic construct"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:32630"
     gene   1..321
                    /gene="ALP-237"
     CDS   1..321
                    /gene="ALP-237"
                    /codon_start=1
                    /transl_table=11
                    /protein_id="CAA00746.1"
                    /db_xref="GI:413357"
                    /translation="SGKSFKAGVCPPKKSAQCLRYKKPEQSDWQCPGKKRCCPDTCG
                    IKCLDPVDTPNPTRRKPGKCPVTYGQCLLLNPPNFCELDGQCKRDLKCCLGFCGKSCV
                    SPVKA"
ORIGIN

Alignment Scores:
Pred. No.:   1.15e-38   Length: 321
Score: 621.00   Matches: 103
Percent Similarity: 99.07%   Conservative: 3
Best Local Similarity: 96.26%   Mismatches: 1
Query Match: 81.50%   Indels: 0
DB: 6   Gaps: 0

US-10-613-105-2 (1-132) x A08110 (1-321)

Qy 26 SerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeuArg 45
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 AGCGGTAAGTCTTTCAAAGCCGGCGTCTGTCCTCCTAAGAAATCTGCTCAGTGCCTGCGT 60

Qy 46 TyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysPro 65
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 TACAAGAAACCCGAGTGCCAGTCTGACTGGCAGTGTCCCGGTAAGAAGCGTTGTTGTCCT 120

Qy 66 AspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArgArgLys 85
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GACACTTGTGGTATCAAATGCCTGGATCCGGTTGACACCCCGAACCCGACGCGTCGTAAA 180

Qy 86 ProGlyLysCysProValThrTyrGlyGlnCysLeuLeuLeuAsnProProAsnPheCys 105
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 CCCGGAAGTGCCCGGTTACCTACGGTCAGTGCCTGCTGCTGAACCCGCCTAACTTCTGC 240

Qy 106 GluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyLysLysGlyLysSer 125
          |||||||||||||||||||||||||||||||||:::||||||||||||||||||||||
Db 241 GAGCTCGATGGCCAGTGCAAACGAGATCTGAAATGCTGCCTGGGTTTCTGCGGTAAAAGC 300

Qy 126 CysValSerProValLysAla 132
          |||||||||||||||||||||
Db 301 TGCGTTAGCCCGGTTAAGGCT 321

RESULT 30
LOCUS   A08105   321 bp   DNA   linear   PAT 28-JUL-1993
DEFINITION Synthetic ALP-gene 230.
ACCESSION   A08105
VERSION   A08105.1 GI:413346
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 321)
AUTHORS   Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
TITLE   Serine protease inhibitor proteins, medicaments containing them,
            DNA sequences coding for these proteins and methods for producing
            these proteins, medicaments and DNA sequences
JOURNAL   Patent: EP 0373335-A 27 20-JUN-1990;
            Gruenenthal GmbH
FEATURES   Location/Qualifiers

Left column (RESULT 31)

```
source          1..321
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
gene            1..321
                /gene="ALP-230"
CDS             1..321
                /gene="ALP-230"
                /codon_start=1
                /transl_table=11
                /protein_id="CAA00741.1"
                /db_xref="GI:413347"
                /translation="SGKSFKAGVCPPKKSAQCLRYKKPECQSDWQCPGKKRCCPDTCG
                IKCLDPVDTPNPTRRKPGKCPVTEGQCLLLNRPNFCELDGQCKRDLKCCLGLCGKSCV
                SPVKA"
ORIGIN

Alignment Scores:
Pred. No.:               3.88e-38           Length:        321
Score:                   614.00             Matches:       102
Percent Similarity:      99.07%             Conservative:    4
Best Local Similarity:   95.33%             Mismatches:      1
Query Match:             80.58%             Indels:          0
DB:                      6                  Gaps:            0

US-10-613-105-2 (1-132) x A08105 (1-321)

Qy   26  SerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeuArg  45
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  AGCGGTAAGTCTTTCAAAGCCGGCGTCTGTCCTCCTCCTCCTAAGAAATCTGCTCAGTGCCTGCGT  60

Qy   46  TyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysPro  65
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  TACAAGAAACCCGAGTGCCAGTCTGACTGGCAGTGTCCTGGTAAGAAGCGTTGTTGTCCT  120

Qy   66  AspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArgArgLys  85
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GACACTTGTGGTATCAAATGCCTGGATCCGGTTGACACCCCGAACCCGACGCGTCGTAAA  180

Qy   86  ProGlyLysCysProValThrTyrGlyGlnCysLeuLeuMetLeuAsnProAsnPheCys  105
         |||||||||||||||||||||||||||::|||||||||||||||||||||||||||||||
Db  181  CCCGGGAAGTGCCCGGTTACCTACGGTCAGTGCCTGCTGATGCTGAACCCGCCTAACTTCTGC  240

Qy  106  GluMetAspGlyGlnCysLysArgAspLeuLysCysCysLeuGlyLeuCysGlyLysSer  125
         ||||||||||||||||::||||||||||||::|||||:::||||||||||||::||||||
Db  241  GAGATGGATGGCCAGTGCAAACGAGATCTGAAATGCTGCCTGGGTCTGTGCGGTAAAAGC  300

Qy  126  CysValSerProValLysAla  132
         |||||||||||||||||||
Db  301  TGCGTTAGCCCGGTTAAGGCT  321
```

RESULT 31
A08109
LOCUS   A08109   321 bp   DNA   linear   PAT 28-JUL-1993
DEFINITION Synthetic ALP-gene 236.
ACCESSION   A08109
VERSION   A08109.1 GI:413354
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 321)
  AUTHORS   Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
  TITLE   Serine protease inhibitor proteins, medicaments containing them,
            DNA sequences coding for these proteins and methods for producing
            these proteins, medicaments and DNA sequences
  JOURNAL   Patent: EP 0373335-A 31 20-JUN-1990;
            Gruenenthal GmbH
FEATURES   Location/Qualifiers
     source   1..321
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"

Right column (RESULT 32)

```
gene            1..321
                /gene="ALP-236"
CDS             1..321
                /gene="ALP-236"
                /codon_start=1
                /transl_table=11
                /protein_id="CAA00745.1"
                /db_xref="GI:413355"
                /translation="SGKSFKAGVCPPKKSAQCLRYKKPECQSDWQCPGKKRCCPDTCG
                IKCLDPVDTPNPTRRKPGKCPVTEGQCLLLNRPNFCELDGQCKRDLKCCLGLCGKSCV
                SPVKA"
ORIGIN

Alignment Scores:
Pred. No.:               3.88e-38           Length:        321
Score:                   614.00             Matches:       102
Percent Similarity:      99.07%             Conservative:    4
Best Local Similarity:   95.33%             Mismatches:      1
Query Match:             80.58%             Indels:          0
DB:                      6                  Gaps:            0

US-10-613-105-2 (1-132) x A08109 (1-321)

Qy   26  SerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeuArg  45
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  AGCGGTAAGTCTTTCAAAGCCGGCGTCTGTCCTCCTCCTCCTAAGAAATCTGCTCAGTGCCTGCGT  60

Qy   46  TyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysPro  65
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  TACAAGAAACCCGAGTGCCAGTCTGACTGGCAGTGTCCTGGTAAGAAGCGGTTGTTGTCCT  120

Qy   66  AspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThrArgArgLys  85
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GACACTTGTGGTATCAAATGCCTGGATCCGGTTGACACCCCGAACCCGACGCGTCGTAAA  180

Qy   86  ProGlyLysCysProValThrTyrGlyGlnCysLeuMetLeuAsnProAsnPheCys  105
         |||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||
Db  181  CCCGGGAAGTGCCCGGTTACCGAAGTCAGTGCCTGCTGCTGTCGTGAACCCGCCTAACTTCTGC  240

Qy  106  GluMetGlyGlyGlnCysLysArgAspLeuLysCysCysMetGlyLeuCysGlyLysSer  125
         |||||||||||||||||||||||||||||:::|||||:::||||||||||||::||||||
Db  241  GAGCTCGATGGCCAGTGCAAACGAGATCTGAAATGCTGCCTGGGTCTGTGCGGTAAAAGC  300

Qy  126  CysValSerProValLysAla  132
         |||||||||||||||||||
Db  301  TGCGTTAGCCCGGTTAAGGCT  321
```

RESULT 32
HSSLIPG
LOCUS   HSSLIPG   2657 bp   DNA   linear   PRI 30-MAR-1995
DEFINITION Human SLPI gene for secretory leukocyte protease inhibitor.
ACCESSION   X04502
VERSION   X04502.1 GI:36485
KEYWORDS   elastase inhibitor; protease inhibitor; secretory leucocyte
            protease inhibitor; trypsin inhibitor.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2657)
  AUTHORS   Stetler,G., Brewer,M.T. and Thompson,R.C.
  TITLE   Isolation and sequence of a human gene encoding a potent inhibitor
            of leukocyte proteases
  JOURNAL   Nucleic Acids Res. 14 (20), 7883-7896 (1986)
   MEDLINE   87040761
   PUBMED   3640338
REFERENCE   2 (bases 1 to 2657)
  AUTHORS   Rogaev,E.I., Keryanov,S.A. and Malyako,Y.K.
  TITLE   Dinucleotide repeat polymorphisms at the P1, HBE1 and MYH7 loci
  JOURNAL   Hum. Mol. Genet. 1 (4), 285 (1992)
   MEDLINE   93265040
   PUBMED   1363870
FEATURES   Location/Qualifiers
```

```
source          1..2657
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
promoter        181..184
                /note="put. CAAT-box"
promoter        242..247
                /note="put. TATA-box"
precursor_RNA   276..2574
                /note="put. primary transcript of SLPI"
mRNA            276..375
                /note="put. exon 1"
CDS             join(291..375,1092..1250,1668..1817,2397..2401)
                /codon_start=1
                /product="secretory leukocyte protease inhibitor (SLPI)"
                /protein_id="CAA28187.1"
                /db_xref="GI:758101"
                /db_xref="GOA:P03973"
                /db_xref="Swiss-Prot:P03973"
                /translation="MKSSGLFPFIVLLALGTLAPWAVEGSGKSFKAGVCPPKKSAQCL
                RYKKPECQSDWQCPGKKRCCPDTCGIKCLDPVDTPNPTRRKPGKCPVTYGQCLMLNPP
                NFCEMDGQCKRDLKCCMGMCGKSCVSPVKA"
sig_peptide     291..365
                /note="signal peptide (AA -25 to -1)"
mat_peptide     join(366..375,1092..1250,1668..1817,2397..2398)
                /gene="SLPI"
                /product="unnamed"
gene            join(366..375,1092..1250,1668..1817,2397..2398)
                /gene="SLPI"
intron          376..1091
                /gene="SLPI"
                /note="intron I"
mRNA            1092..1250
                /gene="SLPI"
                /note="exon 2"
intron          1251..1667
                /gene="SLPI"
                /note="intron II"
mRNA            1668..1817
                /gene="SLPI"
                /note="exon 3"
intron          1818..2396
                /gene="SLPI"
                /note="intron III"
mRNA            2397..2574
                /gene="SLPI"
                /note="exon 4"
misc_feature    2549..2554
                /note="polyA signal"
polyA_site      2574
                /note="polyA site"
ORIGIN

Alignment Scores:
Pred. No.:                              Length:         2657
Score:          3.78e-32                Matches:        109
Percent Similarity:     42.49%          Conservative:   7
Best Local Similarity:  39.93%          Mismatches:     11
Query Match:    71.78%                  Indels:         146
DB:                                     Gaps:           2

US-10-613-105-2 (1-132) x HSSLIPG (1-2657)

Qy     5 GlyLeuPhePheLeuValLeuLeuAlaLeuGlyThrLeu------------- 18
         ||||||||||||||:::||||||      :: :::
Db   998 GGTCTCTCCCCCTCAAAGTGGTGCTTTGTTCAGGAGGCATGATGGTCCTCAGTACCCA 1057

Qy    19 --AlaProTrpAlaValGluGlySerPheLysSerPheLysAlaGlyValCysProPro 37
             :::  ::: :::  ||| |||:::||||||||||||||||
Db  1058 GCCTCCTCCTACCTCTGACTTTCTTCAAAAGCCTTCAAAGCTGGAGTCGTCCTCCT 1117

Qy    38 LysLysSerAlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCys 57
         ||||||||:::||||||||||||||||||||||||||
Db  1118 AAGAAATCTGCCCAGTGCCTTAGATACAAGAAACCTGAGTGCCAGAGTGACTGGCAGTGT 1177
```

```
Qy    58 ProGlyLysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAsp 77
         |||||||||||||||||||||||||||||||||||||
Db  1178 CCAGGGAAGAAGAGATGTTGTCCTGACACTTGTGGCATCAAATGCCTGGATCCTGTTGAC 1237

Qy    78 ThrProAsnProThr---------------------------------------- 82
         |||||||||||:::
Db  1238 ACCCCAAACCAAGTAAGCAAGTCGGGAACTGGCTAGAGAGAGATAGCTGGGGACACAGC 1297

Qy    82 ------------------------------------------------------ 82
         ------------------------------------------------------
Db  1298 ATTAGAGGAGACGGAACTGGGTGATGGGTCCTGCCAGGCCTCCTTGTCAATGCGCTAGTGA 1357

Qy    82 ------------------------------------------------------ 82
         ------------------------------------------------------
Db  1358 GTCACAGTGCCTAAGAAGTAGCCAGTCGGTGAAGCAGCGGGCATTTAGATAGCCAGG 1417

Qy    82 ------------------------------------------------------ 82
         ------------------------------------------------------
Db  1418 TAGTTGGAAGCCTCCACCTAGTCAGCACTGGGCGGTCTGGCCACCTGCCATAATGGGGGCC 1477

Qy    82 ------------------------------------------------------ 82
         ------------------------------------------------------
Db  1478 TGAAGTTCTAGGAGAGCCAGGTGCTATGTTGTTGGGCCGCCCTTAGGGAGAAGGTGGTGGT 1537

Qy    82 ------------------------------------------------------ 82
         ------------------------------------------------------
Db  1538 GATAGAGGTGGGGAGGGGATGATCCCCCTGCTGAAGCTGACGAGGGGCTCACTCTAAA 1597

Qy    82 ------------------------------------------------------ 82
         ------------------------------------------------------
Db  1598 AAGTGGGGATGGGAGGGGTTGTATAAAGTACAAGGCCTCTGACCGGTAGCCTCACTCTCA 1657

Qy    83 --------ArgArgLysProGlyLysCysProValThrTyrGlnCysLeuMet 98
                 |||||||||||||||||||||||||||||||||||||||
Db  1658 CCCAACCCAGCAGGAGGAGAAGCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTTGATG 1717

Qy    99 LeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCys 118
         |||||||||||||||||||||||||||||||||||||
Db  1718 CTTAACCCCCCCAATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGC 1777

Qy   119 MetGlyMetCysGlyLysSerCysValSerProValLys 131
         |||||||||||||||||||||||||||
Db  1778 ATGGGGCATGTGTGGGAAATCCTGCGTTCCCCTGTGAAA 1816
```

```
RESULT 33
I01501
LOCUS       I01501          498 bp ss-DNA          linear   PAT 21-MAY-1993
DEFINITION  Sequence 2 from Patent US 4845076.
ACCESSION   I01501
VERSION     I01501.1  GI:270157
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 498)
  AUTHORS   Heinzel.R., Appelhans,H., Gassen.H.G., Machleidt,W. and
            Seemuller,U.
  TITLE     DNA sequences coding for proteins having the biological activity of
            HUSI-type I inhibitors, biotechnological methods for the
            preparation of said proteins and pharmaceutical compositions
            containing said proteins
  JOURNAL   Patent: US 4845076-A 2 04-JUL-1989;
            Grunenthal GmbH;;
            DE;
FEATURES             Location/Qualifiers
     source          1..498
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:              9.86e-33        Length:         498
```

```
Score:                545.00      Matches:            90
Percent Similarity:   100.00%     Conservative:        0
Best Local Similarity: 100.00%    Mismatches:          0
Query Match:           71.52%     Indels:              0
DB:                               Gaps:                6

US-10-613-105-2 (1-132) x I01501 (1-498)

Qy     43  CysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArg  62
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     25  TGCCTTAGATACAAGAAACCTGAGTGCCAGTGACTGGCAGTGTCCAGGGAAGAAGAGA  84

Qy     63  CysCysProAspThrCysGlyIleLysCysLeuAspProValAspThrProAsnProThr  82
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     85  TGTTGTCCTGACACTTGTGGCATCAAATGCTTGGATCCTGTTGACCCCAAACCCAACA  144

Qy     83  ArgArgLysProGlyLysProValThrTyrGlyGlnCysLeuMetLeuAsnProPro  102
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145  AGGAGGAAGCCTGGGAAGCCGGTCACTTATGGCCAATGTTTGATGCTTAACCCCCCC  204

Qy    103  AsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetGlyMetCys  122
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205  AATTTCTGTGAGATGGATGGCCAGTGCAAGCGTGACTTGAAGTGTTGCATGGGCATGTGT  264

Qy    123  GlyLysSerCysValSerProValLysAla  132
           |||||||||||||||||||||||||||||
Db    265  GGGAAATCCTGCGTTTCCCCTGTGAAAGCT  294
```

```
RESULT 34
HS30OI2/c
LOCUS       HS30OI2          63796 bp      DNA      linear    PRI 07-APR-2001
DEFINITION  Human DNA sequence from clone RP1-30012 on chromosome 20q12-13.12
            Contains ESTs, GSSs and STSs. Contains part of a novel gene and the
            SLPI gene for secretory leukocyte protease inhibitor
            (antileukoproteinase), complete sequence.
ACCESSION   AL035560
VERSION     AL035560.16  GI:11418457
KEYWORDS    HTG; antileukoproteinase; proteinase inhibitor; SLPI.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 63796)
AUTHORS     Ramsay,H.
TITLE       Direct Submission
JOURNAL     Submitted (29-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 29, 2000 this sequence version replaced gi:10198630.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 20, constructed by the Sanger Centre Chromosome 20
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr20
            IMPORTANT: This sequence is not the entire insert of clone
            RP1-30012 It may be shorter because we sequence overlapping
            sections only once, except for a 100 base overlap.
            The true left end of clone RP3-453C12 is at 63637 in this sequence.
            The true right end of clone RP1-172H20 is at 100 in this sequence.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e. phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
```

```
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. RP1-30012 is from the
            library RPCI-1 constructed by the group of Pieter de Jong. For
            further details see
            http://www.chori.org/bacpac/home.htm
VECTOR:     pCYPAC2.

FEATURES             Location/Qualifiers
     source          1..63796
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="q12-13.12"
                     /clone="RP1-30012"
                     /clone_lib="RPCI-1"
     repeat_region   251..554
                     /note="152 copies 2 mer aa 66% conserved"
     repeat_region   253..556
                     /note="76 copies 4 mer aaga 66% conserved"
     repeat_region   975..1059
                     /note="MIR repeat: matches 110. .197 of consensus"
     repeat_region   1143..1674
                     /note="L1ME3 repeat: matches 5592. .6164 of consensus"
     repeat_region   1641..1934
                     /note="L1M4 repeat: matches 4935. .5225 of consensus"
     repeat_region   1939..2038
                     /note="L1M4 repeat: matches 5628. .5728 of consensus"
     repeat_region   2035..2766
                     /note="L1M4 repeat: matches 4215. .4986 of consensus"
     repeat_region   2764..3024
                     /note="L1M4c repeat: matches 1740. .1979 of consensus"
     repeat_region   3025..3316
                     /note="AluSg repeat: matches 1. .293 of consensus"
     repeat_region   3317..4149
                     /note="L1M4c repeat: matches 802. .1740 of consensus"
     gene            join(5925..5593,6243. .6400)
                     /gene="dJ30012.1"
     mRNA            join(5925..5593,6243. .>6400)
                     /gene="dJ30012.1"
                     /product="dJ30012.1 (putative novel transcript)"
                     /note="match: ESTs: Em:AI222267"
                     /evidence=not_experimental
     repeat_region   6104..6223
                     /note="AluSg/x repeat: matches 174. .298 of consensus"
     repeat_region   7809..8022
                     /note="Charlie2 repeat: matches 3505. .3724 of consensus"
     repeat_region   8057..8416
                     /note="MER52C repeat: matches 894. .1278 of consensus"
     repeat_region   8417..9408
                     /note="t1 repeat: matches 2125. .3119 of consensus"
     repeat_region   9340..10314
                     /note="L1M2 repeat: matches 1213. .2572 of consensus"
     repeat_region   10328..10905
                     /note="L1M3e repeat: matches -125. .225 of consensus"
     misc_feature    complement(11008. .11364)
                     /note="match: GSS: Em:AQ492607"
     repeat_region   11229..11533
                     /note="L1M3e repeat: matches -477. .-184 of consensus"
     repeat_region   12038..12289
                     /note="MER1B repeat: matches 65. .337 of consensus"
     repeat_region   12290..12648
                     /note="L1M1 repeat: matches 1207. .1574 of consensus"
     misc_feature    complement(12725. .13141)
                     /note="match: GSS: Em:AQ760198"
     repeat_region   13052..13153
                     /note="MIR repeat: matches 33. .137 of consensus"
     misc_feature    13131..13392
                     /note="match: GSS: Em:AQ081155"
     repeat_region   13218..13397
                     /note="45 copies 4 mer gaaa 66% conserved"
     repeat_region   13928..14216
                     /note="AluSx repeat: matches 1. .310 of consensus"
```

```
repeat_region    14372..14593
                 /note="L1ME2 repeat: matches 5935. .6162 of consensus"
repeat_region    14495. .15168
                 /note="L1ME2 repeat: matches 5. .295 of consensus"
misc_feature     complement(14950. .15573)
                 /note="match: GSS: Em:AQ241185"
misc_feature     15703. .16159
                 /note="match: GSS: Em:AQ017550"
repeat_region    16175. .16190
                 /note="L1M4 repeat: matches 4145. .4159 of consensus"
repeat_region    16191. .16500
                 /note="AluSq repeat: matches 2. .310 of consensus"
repeat_region    16501. .16608
                 /note="L1M4 repeat: matches 4027. .4145 of consensus"
repeat_region    16659. .16839
                 /note="MER97b repeat: matches 2. .179 of consensus"
repeat_region    16847. .17348
                 /note="L1M4 repeat: matches 3473. .4025 of consensus"
repeat_region    17616. .19408
                 /note="L1MB7 repeat: matches 4367. .6167 of consensus"
repeat_region    19499. .19670
                 /note="86 copies 2 mer at 79% conserved"
repeat_region    19502. .19669
                 /note="42 copies 4 mer tata 79% conserved"
repeat_region    19678. .21535
                 /note="L1M4 repeat: matches 67. .1979 of consensus"
repeat_region    21547. .21682
                 /note="Tigger4(Zombi) repeat: matches 1. .137 of consensus"
repeat_region    21690. .22537
                 /note="Tigger4(Zombi) repeat: matches 1872. .2730 of
                 consensus"
repeat_region    22669. .22996
                 /note="L1M4 repeat: matches 3623. .3953 of consensus"
repeat_region    23636. .23896
                 /note="L1M2 repeat: matches 1943. .2572 of consensus"
repeat_region    23825. .26024
                 /note="L1 repeat: matches 2121. .4337 of consensus"
misc_feature     25994. .26445
                 /note="match: GSS: Em:AQ470847"
repeat_region    26025. .26160
                 /note="LTR13 repeat: matches 796. .936 of consensus"
repeat_region    26161. .26235
                 /note="L1 repeat: matches 5112. .5188 of consensus"
misc_feature     26204. .26681
                 /note="match: GSS: Em:AQ624124"
misc_feature     complement(27205. .27381)
                 /note="match: STS: Em:G06121"
gene             complement(27206. .29528)
                 /gene="SLPI"
mRNA             complement(join(27206. .27386,27969. .28118,28542. .28700,
                 29426. .29528))
                 /gene="SLPI"
                 /product="dJ300I2.2 (secretory leukocyte protease
                 inhibitor (antileukoproteinase)"
                 /note="match: cDNAs: Em:U88093 Em:X04470 Em:U73004
                 Em:X04503 Em:M57446 Em:U94341 Em:AF151982 Em:AF178426
                 Em:AF114471 Em:M31216
                 match: ESTs: Em:R71834 Em:AA683520 Em:AA316675 Em:AI742512
                 Em:AA572950 Em:AA564454 Em:AI862145 Em:AI540954
                 Em:AI638119 Em:AI222907 Em:AI743345 Em:AW081599
                 Em:AW264225 Em:AA460433 Em:AI885550 Em:AW238407
                 Em:AI564623 Em:AA541595 Em:AA132992 Em:H65117 Em:AA397776
                 Em:AA993549 Em:AI377093 Em:AA165295 Em:AW190157
                 Em:AA557989 Em:AA164414 Em:H65171 Em:N27733 Em:AW103765
                 Em:AA460434 Em:R49913 Em:AA026192 Em:AW391622 Em:T28664
                 Em:AA026099 Em:AA026497 Em:AA485776 Em:AI247078
                 Em:AI276322 Em:AI367954 Em:AI826892"
                 /evidence=not_experimental
polyA_site       complement(27206)
                 /gene="SLPI"
polyA_signal     complement(27226. .27231)
                 /gene="SLPI"
repeat_region    27229. .27315
```

```
polyA_signal     /note="L2 repeat: matches 2627. .2710 of consensus"
                 complement(27230. .27235)
                 /gene="SLPI"
CDS              complement(join(27382. .27386,27969. .28118,28542. .28700,
                 29426. .29510))
                 /gene="SLPI"
                 /note="match: proteins: Tr:O44397 Sw:P03973 Tr:Q9WUQ4
                 Sw:P97430 Tr:Q9R0Z8 Tr:O95959 Sw:P09412 Sw:P22298
                 Tr:Q9XS44"
                 /codon_start=1
                 /evidence=not_experimental

Alignment Scores:
Pred. No.:                1                          Length:         63796
Score:                  537.50                       Matches:          103
Percent Similarity:      42.51%                      Conservative:       2
Best Local Similarity:   41.70%                      Mismatches:         1
Query Match:             70.54%                      Indels:           141
DB:                     4.9e-30                      Gaps:               9

US-10-613-105-2 (1-132) x HS30012 (1-63796)

Qy      26 SerGlyLysSerPheLysAlaGlyValCysProProLysSerAlaGlnCysLeuArg    45
Db   28710 TCTTCAAAGCCTTCAAAGCTGGAGTCTCCTCCAAGAAATCTGCCCAGTGCCTTAGA 28651

Qy      46 TyrLysLysProGluCysGlnSerAspTrpGlnCysProGlyLysLysArgCysCysPro    65
Db   28650 TACAAGAAACCTGAGTGCCAGATGACTGGCAGTGTCCAGGGAAGAGAGATGTTGTCCT 28591

Qy      66 AspThrCysGlyIleLysLysCysLeuAspProValAspThrProAsnProThr--------    82
Db   28590 GACACTTGTGGCATCAAATGCCTGGATCCTGTTGACACCCCAAGTAAGCAGGTC 28531

Qy      82 -------------------------------------------------------------    82
Db   28530 GGGGAACTGGGTAGAGAGGAGTGAGCCTGGGACACAGCATTAGAGGGATGAACTGGGT 28471

Qy      82 -------------------------------------------------------------    82
Db   28470 GATGGGTCCTGCCAGGCCTCCTTGTCAATCCGTCAGTGAGTCACACTGCCTAAGCAGGA 28411

Qy      82 -------------------------------------------------------------    82
Db   28410 AGGTAGCCAGCAGCTCGTGAAGCAGCGGGCATTAGATAGCCAGGTAGTTGGAAGCCTCC 28351

Qy      82 -------------------------------------------------------------    82
Db   28350 CACCTAGTCAGCACTGGGTGCTGGCCACCTGCATCAATGGGGGGCCTGAAGTTCTCAGGA 28291

Qy      82 -------------------------------------------------------------    82
Db   28290 GAGCCAGGTGCTATGTTTGGGGGCCCCTTAGGGAGAGAGGTGGTGGTGATAGAGGTGGGG 28231

Qy      82 -------------------------------------------------------------    82
Db   28230 AGGGGATGATCCCCCCTGCTGAAGCTGGAGCAGGGGCTCACTCTAAAAAGTGGGGATGGG 28171

Qy      83 ----------------------------------------------------ArgArg    84
Db   28170 AGGGGTTGTATAAAGTACAACAAGGCCTCGACCGGTAGCCGTCACTCTCACCCCAGCAAGGAGG 28111

Qy      85 LysProGlyLyThrProValThrTyrGlyGlnCysLeuMetLeuAsnProAsnPhe    104
Db   28110 AAGCCTGGGAAGTGCCCAGTGACTTATGGCCAATGTTGATGCTTAACCCCCAATTTC 28051

Qy     105 CysGluMetAspGlyGlnCysLysArgAspLeuLysCysMetGlyMetCysGlyLyLys    124
Db   28050 TGTGAGATGGATGGCCAGTGCAAGCGGGACTTGAAGTGTTGCATGGGCGATGTGTGGGAAA 27991

Qy     125 SerCysValSerProValLys    131
Db   27990 TCCTGCGTTCCCCTGTGAAA 27970
```

```
RESULT 35
PIGALP
LOCUS       PIGALP                   600 bp    mRNA    linear   MAM 07-MAR-1995
DEFINITION  Porcine antileukoproteinase mRNA, complete cds.
ACCESSION   M57446
VERSION     M57446.1  GI:164319
KEYWORDS    antileukoproteinase.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 600)
  AUTHORS   Farmer,S.J., Fliss,A.E. and Simmen,R.C.
  TITLE     Complementary DNA cloning and regulation of expression of the
            messenger RNA encoding a pregnancy-associated porcine uterine
            protein related to human antileukoproteinase
  JOURNAL   Mol. Endocrinol. 4 (8), 1095-1104 (1990)
  MEDLINE   91155942
   PUBMED   2293019
COMMENT     Original source text: Porcine uterus. cDNA to mRNA, clone pALP.
FEATURES             Location/Qualifiers
     source          1..600
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /clone="pALP"
                     /sex="female"
                     /tissue_type="uterus"
     gene            1..600
                     /gene="ALP"
     CDS             44..391
                     /gene="ALP"
                     /codon_start=1
                     /product="antileukoproteinase"
                     /protein_id="AAA63346.1"
                     /db_xref="GI:164320"
                     /translation="MAPWAVEGAENALKGGACPPRKIVQCLRYEKPKCTSDWQCPDKK
                     KCCRDTCAIKCLNPVAITNPVKVKPGKCPVVYGQCMMLNPPNHCKTDSQCLGDLKCCK
                     SMCRDTCVCLTPVKA"
     polyA_site      600
                     /gene="ALP"

ORIGIN

Alignment Scores:
Pred. No.:                          6.81e-30        Length:               600
Score:                              508.50          Matches:              87
Percent Similarity:                 78.29%          Conservative:         14
Best Local Similarity:              67.44%          Mismatches:           26
Query Match:                        66.73%          Indels:               2
DB:                                                 Gaps:                 1
```

```
US-10-613-105-2 (1-132) x PIGALP (1-600)

Qy     5 GlyLeuPheProPheLeuValLeuLeuValAlaLeuGly-ThrLeuAlaProTrpAlaValGl 24
           |||||  |||||  ||||  ||||| :::::|||||||||||||
Db     7 GGCCTCTTGCCCTTC--GTGCTTCTTGCCCTGGGAATCCATGGCACCTTGGGCCGTGGA 63

Qy    24 uGlySerGlyLysSerPheLysCysAlaGlyValCysProLysLysSerAlaGlnCysLe 44
           :::  ||| |||| |||||||||  ||| ||:::||||
Db    64 AGGTGCTGAAAATGCTTGAAAGGGGAGCCTGCCCTCCTAGAAAAATTGTCCAGTGCCT 123

Qy    44 uArgTyrLysLysProGluCysProGlyLysSerAspTrpGlnCysGlnSerAspTrpGlnCysLysLysArgCysCy 64
           |||||||||||| |||||| :::|||| :|:::|||||||
Db   124 TAGATATGAGAAACCCAAGTGCACAAGTGACTGGCAGTGTCCAGACAAGAAGAAATGTTG 183

Qy    64 sProAspThrCysGlyIleLysCysLeuAspProValAlaLysProAspThrProAsnProThrArgAr 84
           | |||||||||:|||::||||||  ||:::
Db   184 CCGAGATACTTGCGCAATCAAATGCCTGAACCCTGTTGCTATCACGAACCCAGTTAAGGT 243

Qy    84 gLysProGlyLysCysProValThrTyrGlyGlnCysMetMetLeuAsnProProAsnPh 104
           |||||  ||||||||||||||||||||||||||||||||||||||||||||
Db   244 GAAGCCTGGGAAGTGTCCAGTGGTCTATGGGCCAGTGTATGATGCTCAACCCCCCAATCA 303
```

```
                    104 eCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCysMetCysGlyLeuLy 124
                          :::::
                    304 CTGCAAGACAGCAGCCAGTGCCTGGTGACTTAAAATGCTGCAAGAGCATGTGCGGGAA 363

                    124 sSerCysValSerProValLysAla 132
                          ||:::||||||||||||||||||
                    364 AGTCTGCCTCACCCCTGTGAAAGCC 388
```

```
RESULT 36
AF178426
LOCUS       AF178426                 490 bp    mRNA    linear   ROD 29-AUG-1999
DEFINITION  Rattus norvegicus secretory leukocyte protease inhibitor mRNA,
            complete cds.
ACCESSION   AF178426
VERSION     AF178426.1  GI:5802679
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 490)
  AUTHORS   Song,X., Zeng,L., Jin,W., Thompson,J., Mizel,D.E., Lei,K.,
            Billinghurst,R.C., Poole,A.R. and Wahl,S.M.
  TITLE     Secretory leukocyte protease inhibitor suppresses the inflammation
            and joint damage of bacterial cell wall-induced arthritis
  JOURNAL   J. Exp. Med. 190 (4), 535-542 (1999)
  MEDLINE   99380355
   PUBMED   10049524
REFERENCE   2  (bases 1 to 490)
  AUTHORS   Song,X.-y. and Zeng,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-AUG-1999) Oral infection & Immunity Branch, National
            Institute of Dental & Craniofacial research, 9000 Rockville Pike,
            Bethesda, MD 20892
FEATURES             Location/Qualifiers
     source          1..490
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Lewis"
                     /db_xref="taxon:10116"
     CDS             9..404
                     /note="slpi"
                     /codon_start=1
                     /product="secretory leukocyte protease inhibitor"
                     /protein_id="AAD51758.1"
                     /db_xref="GI:5802680"
                     /translation="MKSCGLFFLMVLLAIGVLAPWSVEGGKNDAIKIGACPARKPAQC
                     LKLEKPECGTDWECPGKQRCCQDTCGFKCLNPVPIRGPVKKKPGRCVKFQGKCLMLNP
                     PNKCQNDGQCDGKYKCCEGMCGKVCLPPV"

ORIGIN

Alignment Scores:
Pred. No.:                          5.11e-28        Length:               490
Score:                              482.50          Matches:              83
Percent Similarity:                 75.00%          Conservative:         16
Best Local Similarity:              62.88%          Mismatches:           30
Query Match:                        63.32%          Indels:               3
DB:                                                 Gaps:                 2
```

```
US-10-613-105-2 (1-132) x AF178426 (1-490)

Qy     1 MetLysSerSerGlyLeuPhePheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
           |||||||| :|||||||  ||||| :::|||||||||||||
Db     9 ATGAAGTCTTGCGGCCTGTTCCTCTCCATGGTCCTCTGGGTACTTGCTGGACCC 68

Qy    21 TrpAlaValGluGlySerGlyLys-----SerPheLysAlaGlyValCysProProLys 38
           ||||||  |||| |:::
Db    69 TGGAGTGTGGGAGGA---GGCAAAATGATGCTATCAAAATTGGACGGCGCTGCCCTGCTGAGA 125

Qy    39 LysSerAlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysPro 58
           |||||||||  ::::| ||||||||||::::|||||
```

DB: 10   Gaps: 3

US-10-613-105-2 (1-132) x AF151982 (1-667)

```
Qy      1  MetLysSerGlyPhePheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro   20
Db     13  ATGAAGTCCAGCGGCCTGTTCCCTCTCATGGTGCTCCTTGCTCTGGGTGTCCTGGCACCC   72

Qy     21  TrpAlaValGlyLysSerGlyLys------SerPheLysAlaGlyLysValCysProLys   38
Db     73  TGGGCTGTGGGAAGGA----GGCAAAAATGATGCTATCAAAATCGGAGCCTGCCCTGCTAGA  129

Qy     39  LysSerAlaGlnCysLysLeuArgTyrLysLysProGlnSerAspTrpGlnCysPro   58
Db    130  AAGCCTGCCCAGTGCCTTAAACGTGAGAAACCAGAGTCGCTGGGGATGTCCCA  189

Qy     59  GlyLysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThr   78
Db    190  GGAAAGCAGAGATGCTGCCAAGATACCTGTGTTTCAAGTGCCTGAATCCTGTTCCCATT  249

Qy     79  ProAsnProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMet   98
Db    250  CCGTGGACCAGTG---AAGAGACCTGGGAGGTGCCTCAAATTTCAAGGAAAATGTCTGATG  306

Qy     99  LeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysArgAspLeuLysCysCys  118
Db    307  CTTAACCCTCCCAATAAGTGCCAGAATGACGGCCAGTGTGATGGCAAATACAAATGTTGT  366

Qy    119  MetGlyMetCysGlyLysSerCysValSerProVal  130
Db    367  GAGGGTATGTGTGGGAAAGTCTGCCTTCCCCAGTG  402
```

RESULT 37
AF151982
LOCUS   AF151982   667 bp   mRNA   linear   ROD 21-JUL-2004
DEFINITION   Rattus norvegicus secretory leukocyte protease inhibitor (SLPI)
                mRNA, complete cds.
ACCESSION   AF151982
VERSION   AF151982.1 GI:4929548
KEYWORDS   .
SOURCE   Rattus norvegicus (Norway rat)
  ORGANISM   Rattus norvegicus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                Rattus.
REFERENCE   1 (bases 1 to 667)
  AUTHORS   Chen,D., Xu,X., Cheon,Y.P., Bagchi,M.K. and Bagchi,I.C.
  TITLE   Estrogen Induces Expression of Secretory Leukocyte Protease
               Inhibitor in Rat Uterus
  JOURNAL   Biol. Reprod. 71 (2), 508-514 (2004)
  PUBMED   15044260
REFERENCE   2 (bases 1 to 667)
  AUTHORS   Chen,D.H., Xu,X.P., Bagchi,M.K. and Bagchi,I.C.
  TITLE   Molecular cloning and spatio-temporal expression of rat secretory
               leukocyte protease inhibitor (SLPI) in the uterus
  JOURNAL   Unpublished
REFERENCE   3 (bases 1 to 667)
  AUTHORS   Chen,D.H., Xu,X.P., Bagchi,M.K. and Bagchi,I.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (17-MAY-1999) Center for Biomedical Research, The
               Population Council, 1230 York Avenue, New York, NY 100021, USA
FEATURES   Location/Qualifiers
     source   1..667
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="SD"
                     /db_xref="taxon:10116"
     gene   1..667
                     /gene="SLPI"
     CDS   13..405
                     /gene="SLPI"
                     /function="serine protease inhibitor"
                     /codon_start=1
                     /product="secretory leukocyte protease inhibitor"
                     /protein_id="AAD34035.1"
                     /db_xref="GI:4929549"
                     /translation="MKSSGLFPLMVLLALGVLAPWSVEGGKNDAIKIGACPARKPAQC
                     LKRKPECSTDWGCPGKQRCCQDTCGFKCLNPVPIRGPVKKPGRCLKFQGKCLMLNPP
                     NKCQNDGQCDGKYKCCEGMCGKVCLPPV"
ORIGIN

RESULT 38
AF421377
LOCUS   AF421377   680 bp   mRNA   linear   ROD 02-OCT-2002
DEFINITION   Rattus norvegicus secretory leukocyte protease inhibitor precursor,
                mRNA, complete cds.
ACCESSION   AF421377
VERSION   AF421377.1 GI:23451267
KEYWORDS   .
SOURCE   Rattus norvegicus (Norway rat)
  ORGANISM   Rattus norvegicus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                Rattus.
REFERENCE   1 (bases 1 to 680)
  AUTHORS   Wang,X., Li,X., Xu,L., Zhan,Y., Yaish-Ohad,S., Barone,F.C. and
               Feuerstein,G.Z.
  TITLE   Secretory Leukocyte Protease Inhibitor (SLPI) Upregulation in
               Cerebral Stroke: Expression of SLPI Protects Brain from Ischemic
               Injury
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 680)
  AUTHORS   Wang,X.
  TITLE   Direct Submission
  JOURNAL   Submitted (19-SEP-2001) Cardiovascular Sciences, DuPont
               Pharmaceuticals, Rt. 141 Experimental Station, E400/318,
               Wilmington, DE 19880, USA
FEATURES   Location/Qualifiers
     source   1..680
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /note="spontaneously hypertensive rats"
     CDS   23..415
                     /note="SLPI"
                     /codon_start=1
                     /product="secretory leukocyte protease inhibitor
                     precursor"
                     /protein_id="AAN32722.1"
                     /db_xref="GI:23451268"
                     /translation="MKSSGLFPLMVLLALGVLAPWSVEGGKNDAIKIGACPARKPAQC
                     LKRKPECSTDWGCPGKQRCCQDTCGFKCLNPVPIRGPVKKPGRCLKFQGKCLMLNPP

Alignment Scores:
Pred. No.:   5.16e-27
Score: 471.00   Length: 667
Percent Similarity: 74.24%   Matches: 84
Best Local Similarity: 63.64%   Conservative: 14
Query Match: 61.81%   Mismatches: 30
                                   Indels: 4

```
Query Match:            60.83%           Indels:          3
DB:                                      Gaps:            2

US-10-613-105-2 (1-132) x AX642233 (1-396)

Qy      1 MetLysSerSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAGTCCTGGCGGCCTTTTACCTTTCACGGTGCTCCTGCTCTGGGGATCCTGGCACCC 60

Qy     21 TrpAlaValGluGlyLysSerGlyLys------SerPheLysAlaGlyValCysProProLys 38
          |||||||||||||||||||     :::       ||||||
Db     61 TGGACTGTGGAAGGA---GGCAAAAATGATGCTATCAAAATCGGAGCCTGCCCTGCTAAA 117

Qy     39 LysSerAlaGlnCysLeuArgTyrLysLysProGlnCysArgAspTrpGlnSerGlnCysPro 58
          ||||||||||||    :::||    :::||||||||||||||||||
Db    118 AAGCCTGCCCAGTGCCTTAAGCTTGAGAAGCCACAATGCCCGTACTGGCAGTGCCCG 177

Qy     59 GlyLysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThr 78
          ||||||||||:::||||      :::||||||||||||||||
Db    178 GGAAAGCAGAGGTGCTGCCAAGATGCTTGCGGTTCCAAGTGCGTGAATCCGTTCCCATT 237

Qy     79 ProAsnProThrArgArgLysProGlyLysCysProValThrTyrGlyLysGlnCysLeuMet 98
          ||||     :::|||||||||||||||     :::||
Db    238 CGCAAACCAGTGTGGAGGAAGCCTGGGAGGTGCGTCAAAACTCAGGCAAGATGATGATG 297

Qy     99 LeuAsnProProAsnPheCysGluMetAspGlyLysMetAspGlyLysArgAspLeuLysCysCys 118
          |||||     :::||||
Db    298 CTTAACCCTCCCAATGTCTGCGAGATGGACGGCAGTGTGACGGCAAATACAACAGTGTGT 357

Qy    119 MetGlyMetCysGlyLysSerCysValSerProVal 130
          |||||||||||||||||||||
Db    358 GAGGGTATATGTGGGAAAGTCTGCCTGCCCCGATG 393

RESULT 40
AX642236/c
LOCUS       AX642236                    396 bp    DNA             linear   PAT 21-FEB-2003
DEFINITION  Sequence 23 from Patent WO0248310.
ACCESSION   AX642236
VERSION     AX642236.1  GI:28474684
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Pittman,D.D., Feldman,J.L., Shields,K.M. and Trepicchio,W.L.
  TITLE     Methods and compositions for diagnosing and treating rheumatoid
            arthritis
  JOURNAL   Patent: WO 0248310-A 23 20-JUN-2002;
            Genetics Instiute, LLC (US)
FEATURES    Location/Qualifiers
     source 1..396
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="Probe"
ORIGIN

Alignment Scores:
Pred. No.:                            Length:           396
Score:          1.12e-26              Matches:           79
Percent Similarity:  72.73%           Conservative:      17
Best Local Similarity: 59.85%         Mismatches:        33
Query Match:         60.83%           Indels:             3
DB:                                   Gaps:               2

US-10-613-105-2 (1-132) x AX642236 (1-396)

Qy      1 MetLysSerSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    396 ATGAAGTCCTGGCGGCCTTTTACCTTTCACGGTGCTCCTGCTCTGGGGATCCTGGCACCC 337

Qy     21 TrpAlaValGluGlyLysSerGlyLys------SerPheLysAlaGlyValCysProProLys 38
          |||||||||||||||||||     :::       ||||||
```

```
      sig_peptide      NKCQNDGQCDGKYKCCEGMCGKVCLPPV"
      3'UTR            23..97
      polyA_site       416..680
ORIGIN                 652

Alignment Scores:
Pred. No.:                            Length:           680
Score:          5.26e-27              Matches:           84
Percent Similarity:  74.24%           Conservative:      14
Best Local Similarity: 63.64%         Mismatches:        30
Query Match:         61.81%           Indels:             4
DB:                                   Gaps:               3

US-10-613-105-2 (1-132) x AF421377 (1-680)

Qy      1 MetLysSerSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||    :::
Db     23 ATGAAGTCCAGCGGCCTGTTCCCTCCTTGCTCCTTGTCTCTGGGTGTCTGGGCACCC 82

Qy     21 TrpAlaValGluGlyLysSerGlyLys------SerPheLysAlaGlyValCysProProLys 38
          ||||||||||||||||||||     :::       ||||
Db     83 TGGAGTGTGGAAGGA---GGCAAAAATGATGCTATCAAAATCGGAGCCTGCCCTGCTAGA 139

Qy     39 LysSerAlaGlnCysLeuArgTyrLysLysProGluCysArgTrpAspTrpGlnSerGlnCysPro 58
          |||||||||||||     :::||||||||||||||
Db    140 AAGCCTGCCCAGTGCCTTAAACGTGAGAAACCAGAGTGCAGTACTGACTGGGATGCCCA 199

Qy     59 GlyLysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThr 78
          ||||||||||:::||      :::||||||||||||||||
Db    200 GGAAAGCAGAGATGCTGCCAAGATACCTGTGGTTTCAAATGCCTGAATCCTGTTCCCATT 259

Qy     79 ProAsnProThrArgArgLysProGlyLysCysProValThrTyrGlyLysGlnCysLeuMet 98
          ||||     :::||||||||||||||     :::||
Db    260 CGTGGACCAGTG---AAGAAGCCTGGGAGGTGCCTCAAATTCAAGGAAAATGTCTGATG 316

Qy     99 LeuAsnProProAsnPheCysGluMetAspGlyLysMetAspGlyLysArgAspLeuLysCysCys 118
          |||||     :::||||
Db    317 CTTAACCCTCCCAATAAGTGCCAGAATGACGGCCAGTGTGATGGCAAATACAACGTTGT 376

Qy    119 MetGlyMetCysGlyLysSerCysValSerProVal 130
          |||||||||||||||||||||
Db    377 GAGGGCATGTGTGGGAAAGTCTGCCTTCCCCCAGTG 412

RESULT 39
AX642233
LOCUS       AX642233                    396 bp    DNA             linear   PAT 21-FEB-2003
DEFINITION  Sequence 20 from Patent WO0248310.
ACCESSION   AX642233
VERSION     AX642233.1  GI:28474681
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Pittman,D.D., Feldman,J.L., Shields,K.M. and Trepicchio,W.L.
  TITLE     Methods and compositions for diagnosing and treating rheumatoid
            arthritis
  JOURNAL   Patent: WO 0248310-A 20 20-JUN-2002;
            Genetics Instiute, LLC (US)
FEATURES    Location/Qualifiers
     source 1..396
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="Probe"
ORIGIN

Alignment Scores:
Pred. No.:                            Length:           396
Score:          1.12e-26              Matches:           79
Percent Similarity:  72.73%           Conservative:      17
Best Local Similarity: 59.85%         Mismatches:        33
```

Db   336   TGGACTGTGGAAGGA---GGCAAAAATGATGCTATCAAAATCGGAGCCTGCCCTGCTAAA 280

Qy   39   LysSerAlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysPro 58

Db   279   AAGCCTGCCCAGTGCCTTAAGCTTGAGAAGCCACAGTGCCGTACTGACTGGGAGTGCCCG 220

Qy   59   GlyLysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThr 78

Db   219   GGAAAGCAGAGGTGTGCCAAGATGCTTGCCGTTCCAAGTGCGTGAATCCTGTTCCCATT 160

Qy   79   ProAsnProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMet 98

Db   159   CGCAAACCAGTGTGGAGGAGAGCCTGGGAGGTGCGTCAAAACTCAGGCAAGATGTATGATG 100

Qy   99   LeuAsnProProAsnPheCysGluMetAspGlyMetCysGlyGlnCysLysArgAspLeuLysArgLeuLysCysCys 118

Db   99   CTTAACCCTCCCAATGTCTGCCAGAGGACGGGCAGTGTGACGGGCAAATACAAGTGCTGT 40

Qy   119   MetGlyMetCysGlyLysSerCysValSerProVal 130

Db   39   GAGGGTATATGTGGGAAAGTCTGCCCGCCCCCGATG 4

RESULT 41
AX577885
LOCUS   AX577885   409 bp   DNA   linear   PAT 08-JAN-2003
DEFINITION   Sequence 7 from Patent WO02081745.
ACCESSION   AX577885
VERSION   AX577885.1   GI:27647093
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Garcia,T., roman Roman,S., Baron,R., Call,K., Theilhaber,J.,
    Connolly,T., Jackson,A., Bushnell,S.E. and Rawadi,G.
  TITLE   Genes involved in osteogenesis, and methods of use
  JOURNAL   Patent: WO 02081745-A 7 17-OCT-2002;
    Aventis Pharma S.A. (FR)
FEATURES   Location/Qualifiers
  source   1..409
    /organism="Mus musculus"
    /mol_type="unassigned DNA"
    /db_xref="taxon:10090"
ORIGIN

Alignment Scores:
Pred. No.:   1.16e-26   Length: 409
Score: 463.50   Matches: 79
Percent Similarity: 72.73%   Conservative: 17
Best Local Similarity: 59.85%   Mismatches: 33
Query Match: 60.83%   Indels: 3
DB: 6   Gaps: 2

US-10-613-105-2 (1-132) x AX577885 (1-409)

Qy   1   MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20

Db   9   ATGAAGTCCTCGGCGCCTTTACCTTTCACGGTGCTCCTTGCCTCTGGGGATCCTGCACCG 68

Qy   21   TrpAlaValGluGlySerGlyLys-----SerPheLysAlaGlyValCysProProLys 38

Db   69   TGGACTGTGGAAGGA---GGCAAAAATGATGCTATCAAAATCGGAGCCTGCCCTGCTAAA 125

Qy   39   LysSerAlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysPro 58

Db   126   AAGCCTGCCCAGTGCCTTAAGCTTGAGAAGCCACAGTGCCGTACTGACTGGGAGTGCCACCG 185

Qy   59   GlyLysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThr 78

Db   186   GGAAAGCAGAGGTGTGCCAAGATGCTTGCCGTTCCAAGTGCGTGAATCCTGTTCCCATT 245

Qy   79   ProAsnProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMet 98

Db   246   CGCAAACCAGTGTGGAGGAGAGCCTGGGAGGTGCGTCAAAACTCAGGCAAGATGTATGATG 305

Qy   99   LeuAsnProProAsnPheCysGluMetAspGlyMetCysGlyGlnCysLysArgAspLeuLysArgLeuLysCysCys 118

Db   306   CTTAACCCTCCCAATGTCTGCCAGAGGACGGGCAGTGTGACGGGCAAATACAAGTGCTGT 365

Qy   119   MetGlyMetCysGlyLysSerCysValSerProVal 130

Db   366   GAGGGTATATGTGGGAAAGTCTGCCCTGCCCCCGATG 401

RESULT 42
MMU88093
LOCUS   MMU88093   671 bp   mRNA   linear   ROD 19-APR-1997
DEFINITION   Mus musculus secretory leukocyte protease inhibitor mRNA, complete
    cds
ACCESSION   U88093
VERSION   U88093.1   GI:1945382
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 671)
  AUTHORS   Zitnik,R.J., Zhang,J., Kashem,M.A., Kohno,T., Lyons,D.E.,
    Wright,C.D., Rosen,E., Goldberg,I. and Hayday,A.C.
  TITLE   The cloning and characterization of a murine secretory leukocyte
    protease inhibitor cDNA
  JOURNAL   Biochem. Biophys. Res. Commun. 232 (3), 687-697 (1997)
  MEDLINE   97271386
  PUBMED   9126337
REFERENCE   2 (bases 1 to 671)
  AUTHORS   Zitnik,R.J., Zhang,J., Kashem,M.A., Kohno,T., Lyons,D.E.,
    Wright,C.D., Rosen,E., Goldberg,I. and Hayday,A.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (31-JAN-1997) Internal Medicine, Yale University, 333
    Cedar Street, New Haven, CT 06520, USA
FEATURES   Location/Qualifiers
  source   1..671
    /organism="Mus musculus"
    /mol_type="mRNA"
    /db_xref="taxon:1090"
    /cell_line="ras-transformed 3T3 cells"
  CDS   13..408
    /codon_start=1
    /function="serine protease inhibitor"
    /product="secretory leukocyte protease inhibitor"
    /protein_id="AAC53140.1"
    /db_xref="GI:1945383"
    /translation="MKSCGLLPFTVLLALGILAPWTVEGGKNDAIKIGACPAKKPAQC
    LKLEKPQCRTDWECPGKQRCCQDACGGKCVNPVPIRKPVWRKPGRCVKTQARCMMLNP
    PNVCQRDGQCDGKYKCCEGICGKVCLPPM"
ORIGIN

Alignment Scores:
Pred. No.:   1.91e-26   Length: 671
Score: 463.50   Matches: 79
Percent Similarity: 72.73%   Conservative: 17
Best Local Similarity: 59.85%   Mismatches: 33
Query Match: 60.83%   Indels: 3
DB: 10   Gaps: 2

US-10-613-105-2 (1-132) x MMU88093 (1-671)

Qy   1   MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20

Db   13   ATGAAGTCCTCGGCGCCTTTACCTTTCACGGTGCTCCTTGCCTCTGGGGATCCTGCACCC 72

Qy   21   TrpAlaValGluGlySerGlyLys-----SerPheLysAlaGlyValCysProProLys 38

Db   73   TGGACTGTGGAAGGA---GGCAAAAATGATGCTATCAAAATCGGAGCCTGCCCTGCTAAA 129

Qy   39   LysSerAlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysPro 58

Db   130 AAGCCTGCCCAGTGCCTTAAGCTGAGAGCCACAATGCCGTACTGACTGGGAGTGCCCCG 189
           ||| :::||||||||||||| :::||||||||| :::|||||||::|||||||
Qy   59 GlyLysLysArgCysProAspThrCysGlyIleLysCysLeuLysAspProValAspThr 78

Db   190 GGAAAGCAGAGGTGCTGCCAAGATGCTTGCGGTTCCAAGATGCCGTGAATCCTGTTCCCATT 249
           ||| :::|||| ||||||::::::|||||:::|||||::|||||||
Qy   79 ProAsnProThrArgArgCysProGlyLysProValThrTyrGlyIleMetCysLeuMet 98

Db   250 CGCAAACCAGTGTGGAGGACGGGACGGGCAGTGTGAAACTCAGGCAAGATGTATGATG 309
           ::::|||||||::|||||||::|||::||||:::|||||||||||::|||||||
Qy   99 LeuAsnProProAsnPheCysGluMetAspGlyArgGlyLysArgAspLeuLysLysCys 118

Db   310 CTTAACCCTCCAATGTCTGCCAGAGGGACGGGCAGTGTGACGGCGCAAATACAAGTGCTGT 369
           ::::|||||||::|||
Qy   119 MetGlyMetCysGlyLysSerCysValSerProVal 130

Db   370 GAGGGTATATGTGGGAAAGTCTGCCTGCCCCCGATG 405
           |||::|::::
RESULT 43
MMU94341
MMU94341
LOCUS   MMU94341   682 bp   mRNA   linear   ROD 29-OCT-1997
DEFINITION   Mus musculus secretory leukoprotease inhibitor mRNA, complete cds.
ACCESSION   U94341
VERSION   U94341.1 GI:1945450
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 682)
  AUTHORS   Abe,T., Tominaga,Y., Kikuchi,T., Watanabe,A., Satoh,K., Watanabe,Y.
             and Nukiwa,T.
  TITLE   Bacterial pneumonia causes augmented expression of the secretory
             leukoprotease inhibitor gene in the murine lung
  JOURNAL   Am. J. Respir. Crit. Care Med. 156 (4 Pt 1), 1235-1240 (1997)
  MEDLINE   98011992
  PUBMED   9351627
REFERENCE   2 (bases 1 to 682)
  AUTHORS   Abe,T., Tominaga,Y., Kikuchi,T., Watanabe,A., Satoh,K., Watanabe,Y.
             and Nukiwa,T.
  TITLE   Direct Submission
  JOURNAL   Submitted (18-MAR-1997) Department of Respiratory Oncology and
             Molecular Medicine, Institute of Development, Aging and Cancer,
             Tohoku University, 4-1 Seiryomachi, Aoba-ku, Sendai 980-77, Japan
FEATURES   Location/Qualifiers
     source   1..682
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="B6CBA [C57Black/6 x CBA]"
                     /db_xref="taxon:10090"
                     /tissue_type="lung"
     CDS   14..409
                     /codon_start=1
                     /product="secretory leukoprotease inhibitor"
                     /protein_id="AAC53394.1"
                     /db_xref="GI:1945451"
                     /translation="MKSCGLLPFTVLLALGILAPWTVEGGKNDAIKIGACPAKKPAQC
                     LKLEKPQCRTDWECPGKQRCCQDACGSKCVNPVPIRKPVWRKPGRCVKTQARCMMLNP
                     PNVCQRDGQCDGKYKCCEGICGKVCLPPM"
ORIGIN

Alignment Scores:
Pred. No.:   682
Score: 1.94e-26
Percent Similarity: 463.50   Length: 682
Best Local Similarity: 72.73%   Matches: 79
Query Match: 59.85%   Conservative: 17
DB: 60.83%   Mismatches: 33
                                              Indels: 3
                                              Gaps: 2

US-10-613-105-2 (1-132) x MMU94341 (1-682)

Qy   1 MetLysSerSerGlyLeuPheProPheLeuValLeuAlaLeuGlyThrLeuAlaPro 20

Db   130 AAGAAGCTCTCGGCCTTTACCTTTCACGGTGCTCCTTGCTCTCTGGGGATCCTGGCACCC 73
           |||||||| ||||||||
Qy   Db   14 ATGAAGTCCTGCGGCCTTTTACCTTTCACGGTGCTCCTTGCTCTCTGGGGATCCTGGCACCC 73

Db   21 TrpAlaValGluGlySerGlyLys------SerPheLysAlaGlyValCysProProLys 38
           :::
Db   74 TGGACTGTGGAAGGA---GGCAAAATGATGTCATCAAAATCGGAGCTCCGCTGCTGTAAA 130

Qy   39 LysSerAlaGlnCysLeuArgTyrLysProGluCysTrpGlnSerAspTrpGlnCysPro 58
Db   131 AAGCCTGCCCAGTGCCTTAGGCTTGAGAAGCCACAATGCCGTACTGACTGGGAGTGCCCG 190

Qy   59 GlyLysLysArgCysProAspThrCysGlyIleLysCysLeuLysAspProValAspThr 78
Db   191 GGAAAGCAGAGGTGCTGCCAAGATGCTTGCGGTTCCAAGATGCCGTGAATCCTGTTCCCATT 250

Qy   79 ProAsnProThrArgArgCysProGlyLysProValThrTyrGlyIleMetCysLeuMet 98
Db   251 CGCAAACCAGTGTGGAGGACGGGACGGGCAGTGTGAAACTCAGGCAAGATGTATGATG 310

Qy   99 LeuAsnProProAsnPheCysGluMetAspGlyArgGlyLysArgAspLeuLysLysCys 118
Db   311 CTTAACCCTCCAATGTCTGCCAGAGGGACGGGCAGTGTGACGGCGCAAATACAAGTGCTGT 370

Qy   119 MetGlyMetCysGlyLysSerCysValSerProVal 130
Db   371 GAGGGTATATGTGGGAAAGTCTGCCTGCCCCCGATG 406

RESULT 44
BC028509
LOCUS   BC028509   894 bp   mRNA   linear   ROD 30-JUN-2004
DEFINITION   Mus musculus secretory leukocyte protease inhibitor, mRNA (cDNA
             clone MGC:41142 IMAGE:1513866), complete cds.
ACCESSION   BC028509
VERSION   BC028509.1 GI:20306995
KEYWORDS   MGC.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 894)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
             Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
             Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
             Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
             Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
             Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
             Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
             Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
             Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
             McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
             Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
             Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
             Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
             Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
             Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
             Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
             Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
             Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE   Generation and initial analysis of more than 15,000 full-length
             human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  PUBMED   12477932
REFERENCE   2 (bases 1 to 894)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (23-APR-2002) National Institutes of Health, Mammalian
             Gene Collection (MGC), Cancer Genomics Office, National Cancer
             Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
             USA
  REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
  COMMENT   Contact: MGC help desk
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Marcello Bento Soares, Ph.D.

cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 67 Row: j Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6755573.

```
FEATURES         Location/Qualifiers
     source          1..894
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="MGC:41142 IMAGE:1513866"
                     /tissue_type="Mammary gland, lactating mouse"
                     /clone_lib="Soares_mammary_gland_NMLMG"
                     /lab_host="DH10B"
                     /note="vector: pT7T3-Pac"
     gene            1..894
                     /gene="slpi"
                     /db_xref="LocusID:20568"
                     /db_xref="MGI:109297"
     CDS             236..631
                     /gene="slpi"
                     /codon_start=1
                     /product="secretory leukocyte protease inhibitor"
                     /protein_id="AAH28509.1"
                     /db_xref="GI:20306996"
                     /db_xref="LocusID:20568"
                     /db_xref="MGI:109297"
                     /translation="MKSCGLLPFTVLLALGILAPWTVEGGKNDAIKIGACPAKKPAQC
                     LKLEKPQCRTDWECPGKWRCCQDACGSKCVNPVPIRKPVWRKPGRCVKTQARCMMLNP
                     PNVCQRDGQCDGKYKCCEGICGKVCLPPM"
ORIGIN
```

Alignment Scores:
Pred. No.:   2.5e-26   Length: 894
Score: 463.50   Matches: 79
Percent Similarity: 72.73%   Conservative: 17
Best Local Similarity: 59.85%   Mismatches: 33
Query Match: 60.83%   Indels: 3
DB: 10   Gaps: 2

US-10-613-105-2 (1-132) x BC028509 (1-894)

```
Qy    1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
          ||||||||||||||||||||||         ||||||||||||||
Db  236 ATGAAGTCCTGCGCCCTTTTACCTTTCACGGTGCTCCTTGCTCTGGGGATCCTGGCACCC 295

Qy   21 TrpAlaValGluGlySerGlyLys------SerPheLysAlaGlyValCysProProLys 38
          ||||||||      ||| :::   |||
Db  296 TGGACTGTGGAAGGA--GGCAAAAATGATGCTATCAAAATCGGAGCCTGCCCTGCTAAA 352

Qy   39 LysSerAlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysPro 58
          |||||||||||| ||||     |||||||||||:::|
Db  353 AAGCCTGCCCAGTGCCTTAAGCTGGAGAAGCCACAATGCCGTACTGACTGGGAGTGCCCG 412

Qy   59 GlyLysTrpArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThr 78
          |||||:||||||||||||||  |||||||||:::::|
Db  413 GGAAAGCAGAGGACGTGCTGCCAAGATGCTTGCGGTTCCAAGTGGGTGAATCCTGTTCCATT 472

Qy   79 ProAsnProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMet 98
          |||  |||::::|||||||:::||||||||||:::  :|||
```

Db 473 CGCAAACCAGTGTGGAGGAAGCCGGAGGTCGTCGTCAAAACTCAGGCAAGATGTATGATG 532
          ||| :::||||||||||||||::::|||
Qy   99 LeuAsnProProAsnPheCysGluMetAspGlyGlnCysLysLysArgAspLysCysCys 118
           |||||||||||||||||:::::
Db 533 CTTAACCCTCCAAATGTCTGCCAGAGGACGACGGGCAGTGTGACGGCAAATACAAGTGTGT 592

Qy 119 MetGlyMetCysGlyLysSerGlyValSerProVal 130
          ||::||||||||||||||||||||
Db 593 GAGGGTATATGTGGGAAAGTCTGCCTGCCCCCGATG 628

RESULT 45
MMU73004

```
LOCUS       MMU73004                1123 bp    mRNA    linear   ROD 27-FEB-1997
DEFINITION  Mus musculus secretory leukocyte protease inhibitor mRNA, complete
            cds.
ACCESSION   U73004
VERSION     U73004.1  GI:1763262
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1123)
  AUTHORS   Jin,F.Y., Nathan,C., Radzioch,D. and Ding,A.
  TITLE     Secretory leukocyte protease inhibitor: a macrophage product
            induced by and antagonistic to bacterial lipopolysaccharide
  JOURNAL   Cell 88 (3), 417-426 (1997)
  MEDLINE   97191310
   PUBMED   9039268
REFERENCE   2  (bases 1 to 1123)
  AUTHORS   Ding,A., Jin,F.-Y. and Nathan,C.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-1996) Medicine, Cornell University Medical
            College, 1300 York Ave. Box 57, New York, NY 10021, USA
FEATURES         Location/Qualifiers
     source          1..1123
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /cell_line="RAW 264.7"
                     /cell_type="macrophage"
     CDS             447..842
                     /codon_start=1
                     /product="secretory leukocyte protease inhibitor"
                     /protein_id="AAC53047.1"
                     /db_xref="GI:1763263"
                     /translation="MKSCGLLPFTVLLALGILAPWTVEGGKNDAIKIGACPAKKPAQC
                     LKLEKPQCRTDWECPGKWRCCQDACGSKCVNPVPIRKPVWRKPGRCVKTQARCMMLNP
                     PNVCQRDGQCDGKYKCCEGICGKVCLPPM"
ORIGIN
```

Alignment Scores:
Pred. No.:   3.22e-26   Length: 1123
Score: 463.50   Matches: 79
Percent Similarity: 72.73%   Conservative: 17
Best Local Similarity: 59.85%   Mismatches: 33
Query Match: 60.83%   Indels: 3
DB: 10   Gaps: 2

US-10-613-105-2 (1-132) x MMU73004 (1-1123)

```
Qy    1 MetLysSerSerGlyLeuPheProPheLeuValLeuLeuAlaLeuGlyThrLeuAlaPro 20
          ||||||||||||||||||||||         ||||||||||||||
Db  447 ATGAAGTCCTGCGGCCCTTTTACCTTTCACGGTGCTCCTTGCTCTGGGGATCCTGGCACCC 506

Qy   21 TrpAlaValGluGlySerGlyLys------SerPheLysAlaGlyValCysProProLys 38
          ||||||||      ||| :::   |||
Db  507 TGGACTGTGGAAGGA--GGCAAAAATGATGCTATCAAAATCGGAGCCTGCCCTGCTAAA 563

Qy   39 LysSerAlaGlnCysLeuArgTyrLysLysProGluCysGlnSerAspTrpGlnCysPro 58
          ||| |||||||| ||||     :::|||||||||:::::|
```

```
Db  564 AAGCCTGCCCAGTGCCTTAAGCTTGAGAAGCCACAATGCCCTACTGACTGGGAGTGCCCG 623

Qy   59 GlyLysLysArgCysCysProAspThrCysGlyIleLysCysLeuAspProValAspThr 78
             |||||::::|||||||||  |||   |||||||||||:::|||||||||||||
Db  624 GGAAAGCAGAGGTGCTGCCAAGATGCTTGCGGTTCCAAGTGCGTGAATCCTGTTCCCATT 683

Qy   79 ProAsnProThrArgArgLysProGlyLysCysProValThrTyrGlyGlnCysLeuMet 98
             |||  |||||||||||||||||:::|||   |||  :::|||:::|||||
Db  684 CGCAAACCAGTGTGGAGGAAGCCTGGGAGGTGCGTCAAAACTCAGGCAAGATGTATGATG 743

Qy   99 LeuAsnProProAsnPheCysGluMetAspGlyCysLysArgAspLeuLysCysCys 118
             ||||||||||||||  :::|||  |||||:::|||||||||
Db  744 CTTAACCCTCCCAAATGTCTGCCAGAGGGACCGGCAGTGTGACGGGCAAATACAAGTGCTGT 803

Qy  119 MetGlyMetCysGlyLysSerCysValSerProVal 130
             |||:::|||||||||:::|||||  |||:::
Db  804 GAGGGTATATGTGGGAAAGTCTGCCTGCCCCCGATG 839
```

Search completed: October 24, 2004, 05:07:16
Job time : 3582 secs